

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:16:09 / Search time 25.3023 Seconds  
(without alignments)  
2067.627 Million cell updates/sec

Title: US-10-068-426-1

Perfect score: 2906

Sequence: 1 MPLLALLLPSPHPHPIC.....MHEALHHYTKSLSPGK 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1693	58.3	626	1 NBHUIA	platelet glycoprotein
2	1205	41.5	374	2 S69339	Ig heavy chain V r
3	1197.5	41.2	330	1 GHU	Ig gamma-1 chain C
4	1197	41.2	255	4 S1856	Ig gamma-1 chain C
5	1150	39.6	234	2 P0207	Ig gamma chain C r
6	1149.5	39.6	326	1 G2HU	Ig gamma-2 chain C
7	1141.5	39.3	377	2 A23511	Ig gamma-3 chain C
8	1139.5	39.2	377	2 A60764	Ig gamma-3 chain C
9	1137	39.1	327	1 G4HU	Ig gamma-4 chain C
10	1115	38.4	289	1 G3HWT	Ig gamma-3 heavy c
11	918	31.6	328	2 I47160	Ig gamma 2b chain
12	918	31.6	328	2 I47159	Ig gamma 2a chain
13	913	31.4	323	1 GHRB	Ig gamma chain C r
14	911	31.3	277	2 I47162	Ig gamma 4 chain c
15	888.5	30.6	328	2 I47158	Ig gamma 1 chain c
16	887	30.5	328	1 G2GP	Ig gamma 2 chain c
17	886	30.5	328	2 I47161	Ig gamma 3 chain c
18	851	29.3	470	2 S22080	Ig heavy chain pre
19	843	29.0	329	1 G3MSC	Ig gamma-3 chain C
20	835	28.7	308	2 C30554	Ig heavy chain C r
21	835	28.7	472	2 S31459	Ig gamma-1 chain -
22	834	28.7	444	2 PC4436	monoclonal antibod
23	832	28.6	398	1 G3MSM	Ig gamma-3 chain C
24	831	28.6	326	2 PS6017	Ig gamma-1 chain C
25	828	28.5	333	2 PS0018	Ig gamma-2c chain
26	823	28.3	329	2 S00847	Ig gamma-2c chain
27	820.5	28.2	324	1 G1MS	Ig gamma-1 chain C
28	815.5	28.1	393	1 G1MSM	Ig gamma-1 chain C
29	812	27.9	322	2 PS0019	Ig gamma-2a chain

30	805	27.7	330	1 G2MSA	Ig gamma-2a chain
31	805	27.7	469	2 S37483	Ig gamma-2a chain
32	802	27.6	335	1 G2MSAB	Ig gamma-2a chain
33	800	27.5	399	1 G2MSAM	Ig gamma-2a chain
34	790	27.2	446	2 S40295	Ig gamma-2a chain
35	769.5	26.5	474	1 G2MS11	Ig gamma-2b chain
36	762.5	26.2	327	2 S06611	Ig gamma-2 chain C
37	762.5	26.2	327	2 S06611	Ig gamma-2 chain C
38	746	25.7	475	2 S01321	Ig gamma-2b chain
39	704	24.2	180	2 I46732	Ig gamma heavy cha
40	580.5	20.0	249	2 S69340	Ig heavy chain VHI
41	573	19.7	152	2 S14236	Ig gamma-1 chain C
42	568.5	19.6	218	2 A36040	Ig heavy chain V-I
43	409	14.1	572	2 B46529	Ig y heavy chain (
44	359.5	12.4	388	1 EHMS	Ig epsilon chain C
45	358	12.3	453	2 S37768	Ig mu chain C regi

ALIGNMENTS

RESULT 1

NBHU1A  
platelet glycoprotein Ib alpha chain precursor - human  
N:Alternate names: membrane glycoprotein Ib alpha chain  
N:Contains: glycoocalicin  
C:Species: Homo sapiens (man)  
C>Date: 28-Dec-1987 #sequence revision 28-Dec-1987 #text change 22-Jun-1999  
C:Accession: A94174; A60435; A94173; S16945; IS5355; A27075; A27102  
R:Lopez, J.A.; Chung, D.W.; Fujikawa, K.; Hagen, F.S.; Papayannopoulou, T.; Roth, G.J  
Proc. Natl. Acad. Sci. U.S.A. 84, 5615-5619, 1987  
A>Title: Cloning of the alpha-chain of human platelet glycoprotein Ib: a transmembran  
A:Reference number: A94174; MUID:87289655; PMID:3303030  
A:Accession: A94174  
A:Molecule type: mRNA  
A:Residues: 1-826 <LOP>  
A:Cross-references: GB:J02940; NID:gl83499; PIDN:AA5595.1; PID:9306793  
R:Wicki, A.N.; Walz, A.; Gerber-Huber, S.N.; Wenger, R.H.; Vornhagen, R.; Clemenston,  
Thromb. Haemost. 61, 448-453, 1989  
A>Title: Isolation and characterization of human blood platelet mRNA and construction  
d cloning of a GPIb coding cDNA insert  
A:Reference number: A60435; MUID:90020160; PMID:2799758  
A:Accession: A60435  
A:Molecule type: mRNA  
A:Residues: 207-467 <WIC>  
R:Titani, K.; Takio, K.; Handa, M.; Ruggeri, Z.M.  
Proc. Natl. Acad. Sci. U.S.A. 84, 5610-5614, 1987  
A>Title: Amino acid sequence of the von Willebrand factor-binding domain of platelet  
A:Reference number: A94173; MUID:87289654; PMID:3497398  
A:Accession: A94173  
A:Molecule type: protein  
A:Residues: 17-315 <TIT>  
R:Hess, D.; Schaller, J.; Rickli, E.E.; Clemenston, K.J.  
Eur. J. Biochem. 199, 389-393, 1991  
A>Title: Identification of the disulphide bonds in human platelet glycoocalicin.  
A:Reference number: S16945; MUID:91301149; PMID:2070794  
A:Accession: S16945  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 224-227;262-270;277-282 <HES>  
R:Lopez, J.A.; Ludwig, E.H.; McCarthy, B.J.  
J. Biol. Chem. 267, 10055-10061, 1992  
A>Title: Polymorphism of human glycoprotein Ib alpha results from a variable number o  
ations.  
A:Reference number: IS5355; MUID:92250564; PMID:1577776  
A:Accession: IS5355  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 412-427 <RES>  
A:Cross-references: GB:S34436; NID:g249176; PIDN:AA22152.1; PID:g249177  
A>Note: variant D  
C:Comment: Glycoprotein Ib (GPIb), a surface membrane protein of platelets, participa  
C:Comment: Platelet activation apparently involves disruption of the macromolecular c

C;Comment: Binding sites for von Willebrand factor and thrombin (the latter site with ur  
C;Comment: Glycocalicin, which is approximately coextensive with the extracellular part  
C;Genetics:  
A;Gene: GDB:GPIBA; GPIB  
A;Cross-references: GDB:118806; OMIM:231200  
A;Map position: 17pter-17p12  
C;Complex: heterodimer with platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein  
C;Superfamily: platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein  
C;Keywords: blood coagulation; duplication; glycoprotein; platelet membrane; tandem repeat  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-626/Product: platelet glycoprotein Ib alpha chain #status predicted <MP>  
F:48-71/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F:72-93/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F:94-116/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:117-140/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F:141-164/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F:165-188/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F:189-212/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F:379-430/Region: proline/threonine-rich 9-residue repeats  
F:502-540/Domain: transmembrane #status predicted <TRM>  
F:541-626/Domain: intracellular #status predicted <INT>  
F:37,175/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:308/Binding site: carbohydrate (Thr) (covalent) #status experimental

Query Match 58.3%; Score 1693; DB 1; Length 626;  
Best Local Similarity 87.4%; Pred. No. 2.9e-108;  
Matches 333; Conservative 4; Mismatches 18; Indels 26; Gaps 5;

Qy 1 MPLLILLLLPSLPHPICEVSVASHLEVNCDKRLTALPDPKDTILHLSNLLY 60  
Db 1 MPLLILLLLPSLPHPICEVSVASHLEVNCDKRLTALPDPKDTILHLSNLLY 60

Qy 61 TFSLATLMPYTRLTQMLDRCLETKLVQDGLPVLGTLDSLHQLSLPLGLGTLPALTV 120  
Db 61 TFSLATLMPYTRLTQMLDRCLETKLVQDGLPVLGTLDSLHQLSLPLGLGTLPALTV 120

Qy 121 LDVSFNRLTSLPGALRGLOELQYLYKGNELKLTLPGLLTPTPKLEKSLANNLTLP 180  
Db 121 LDVSFNRLTSLPGALRGLOELQYLYKGNELKLTLPGLLTPTPKLEKSLANNLTLP 180

Qy 181 AGLNGLENLDTLLQNSLYTIIPKFGFSGHLLPFAFLHGNPMLCNCEILYFRWLQDNA 240  
Db 181 AGLNGLENLDTLLQNSLYTIIPKFGFSGHLLPFAFLHGNPMLCNCEILYFRWLQDNA 240

Qy 241 ENVVYVWQGVVDVKAWSNVASVQCDNSDKPPVYKPGCGPTLGDGDTLDLYYFEEDT 300  
Db 241 ENVVYVWQGVVDVKAWSNVASVQCDNSDKPPVYKPGCGPTLGDGDTLDLYYFEEDT 300

Qy 301 EGDKVRATRTVVKPPTKARPHTCP-----PCPA---PEALGAPSVFLFPPK 343  
Db 301 EGDKVRATRTVVKPPTKARPHTCP-----PCPA---PEALGAPSVFLFPPK 343

Qy 344 --PKDTI-----MISRTPEVT 357  
Db 359 WTPNFTLMEISITSPKPKST 379

RESULT 2  
S69339  
Ig heavy chain V region precursor - human  
C;Species: Homo sapiens (man)  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 01-Dec-2000  
C;Accession: S69339; S72664  
R;Khamilichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.  
Eur. J. Biochem. 229, 54-60, 1995  
A;Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.  
A;Reference number: S69339; MUID:95262687; PMID:7744049  
A;Accession: S69339  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-374 <KHA>  
A;Cross-references: EMBL:X81695  
R;Khamilichi, A.A.

submitted to the EMBL Data Library, September 1994  
A;Reference number: S72664  
A;Accession: S72664  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-140,'C',142-374 <KH2>  
A;Cross-references: EMBL:X81695  
C;Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 41.5%; Score 1205; DB 2; Length 374;  
Best Local Similarity 86.2%; Pred. NO. 4e-75;  
Matches 231; Conservative 5; Mismatches 24; Indels 8; Gaps 3;

Qy 295 DEGDLDLY--DYDEEDTEGKV---RATRTVVKPPTKA--RHTCPCPAPALGAPS 336  
Db 107 DPADTATTCYGYGEGYCGGRFHSWGGTLTVTSSEPKSCDKTHCTCPAPALLGSPS 166

Qy 337 VLFPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTPREEQYNST 396  
Db 167 VLFPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTPREEQYNST 226

Qy 397 TRVSVLTVLHODMLNGEYKCKVSNKALPVPDKITISKAKGQPREPOVYTLPPSRREMT 456  
Db 227 YRWSVLTVLHODMLNGEYKCKVSNKALPVPDKITISKAKGQPREPOVYTLPPSRREMT 286

Qy 457 KNQVSLTCLVKGFPSPDIATVWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQ 516  
Db 287 KNQVSLTCLVKGFPSPDIATVWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQ 346

Qy 517 GNVFSCSVMEALHNHYTKQSLSPGK 544  
Db 347 GNVFSCSVMEALHNHYTKQSLSPGK 374

RESULT 3  
GHU  
Ig gamma-1 chain C region - human  
C;Species: Homo sapiens (man)  
C;Date: 31-Jan-1981 #sequence\_revision 18-Aug-1982 #text\_change 16-Jul-1999  
C;Accession: A93433; S3887; S3887; B90563; A90564; B91668; A91723; A02146  
R;Ellison, J.W.; Berson, B.J.; Hood, L.E.  
Nucleic Acids Res. 10, 4071-4079, 1982  
A;Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.  
A;Reference number: A93433; MUID:82274238; PMID:6287432  
A;Accession: A93433  
A;Molecule type: DNA  
A;Residues: 1-330 <ELL>  
A;Cross-references: EMBL:Z17370  
A;Note: this sequence has the Gln(17) allotypic marker, 97-Lys, and the Gln(11) marker  
A;Note: Lys-330 is removed after translation  
R;Harris, L.J.  
submitted to the EMBL Data Library, October 1992  
A;Reference number: S33904  
A;Accession: S36861  
A;Molecule type: DNA  
A;Residues: 2-330 <HAR>  
A;Cross-references: EMBL:Z17370  
R;Takahashi, N.; Ueda, S.; Obata, M.; Nakai, T.; Honjo, T.  
Cell 29, 671-679, 1982  
A;Title: Structure of human immunoglobulin gamma genes: implications for evolution of  
A;Reference number: S33887; MUID:83001943; PMID:6811139  
A;Accession: S33887  
A;Molecule type: DNA  
A;Residues: 88-113;235-330 <TAK>  
A;Cross-references: EMBL:Z17370  
R;Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman  
Biochemistry 9, 3161-3170, 1970  
A;Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid seq  
A;Reference number: A90563; MUID:71064024; PMID:5489771  
A;Contents: myeloma protein Bu  
A;Accession: B90563  
A;Molecule type: protein  
A;Residues: 1-96,'R',98-135 <CUN>

A:Note: this sequence has the Gln(3) marker, 97-Arg  
 R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.  
 Biochemistry 9, 3171-3181, 1970  
 A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid sequence  
 A:Reference number: A90564; MUID:71064025; PMID:5530842  
 A:Contents: Eu  
 A:Accession: A90564  
 A:Molecule type: protein  
 A:Residues: 136-154, 'Q', 156-165, 'Q', 167-176, 'Q', 178-194, 'N', 196-197, 'D', 199-238, 'E', 240,  
 A:Note: this sequence has the Gln(non-1) markers, 239-Glu and 241-Met  
 R:Ponstingl, H.; Hilschmann, N.  
 Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976  
 A:Title: Die Primärstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nie),  
 igen Primärstruktur.  
 A:Reference number: A91668; MUID:77070269; PMID:826475  
 A:Accession: B91668  
 A:Contents: myeloma protein Nie  
 A:Molecule type: protein  
 A:Residues: 1-34, 'Q', 36-96, 'K', 98-115, 'Q', 117-197, 'D', 199-238, 'D', 240, 'L', 242-268, 'E', 27  
 A:Note: this sequence has the Gln(17) and Gln(1) markers  
 R:Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.  
 Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983  
 A:Title: Die Primärstruktur des kristallisierten monoklonalen Immunglobulins IgG1 KOL  
 A:Reference number: A91723; MUID:83289331; PMID:6884994  
 A:Contents: myeloma protein KOL; disulfide bonds  
 A:Accession: A91723  
 A:Molecule type: protein  
 A:Residues: 1-96, 'R', 98-197, 'D', 199-238, 'E', 240, 'M', 242-266, 'D', 268-271, 'D', 273-330 <SCH  
 A:Note: this sequence has the Gln(3) and Gln(non-1) markers  
 R:Gall, W.E.; Edelman, G.M., 1970  
 Biochemistry 9, 3188-3196, 1970  
 A:Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid  
 A:Reference number: A90565; MUID:71064027; PMID:4923144  
 A:Contents: annotation; disulfide bonds  
 R:Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.  
 Hoppe-Seyler's Z. Physiol. Chem. 357, 1513-1540, 1976  
 A:Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob  
 enamide cleavage products, and the disulfide bridges.  
 A:Reference number: A91667; MUID:77070267; PMID:1002129  
 A:Contents: annotation; disulfide bonds  
 C:Genetics:  
 A:Gene: GDB:IGHG1  
 A:Cross-references: GDB:120085; OMIM:147100  
 A:Map position: 14Q32.33-14Q32.33  
 A:Introns: 99/1; 114/1; 224/1  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:120-85/Domain: immunoglobulin homology <IM1>  
 F:137-206/Domain: immunoglobulin homology <IM2>  
 F:243-310/Domain: immunoglobulin homology <IM3>  
 F:27-83,144-204,250-308/Disulfide bonds: #status experimental  
 F:103/Disulfide bonds: interchain (to light chain) #status experimental  
 F:109,112/Disulfide bonds: interchain (to heavy chain) #status experimental  
 F:180/Binding site: carbohydrate (Aan) (covalent) #status experimental  
 Query Match 41.2%; Score 1197.5; DB 1; Length 330;  
 Best Local Similarity 73.1%; Pred. No. 1.1e-74;  
 Matches 236; Conservative 11; Mismatches 37; Indels 39; Gaps 6;  
 QY 253 KAMTSNVAQCDNSDKFP-----VYKPGKCPGLGDEGDTLDYVYPE 297  
 DB 16 KSTSGTAALGCLVXDYFPEPVTVSNWSGALTSVHTFPA----VLQSSG---LYSLSSV 68  
 QY 298 ETEGDKVATRTVV-----KFTKARP-----HTCPCPAPALGAPSVFLFP 341  
 DB 69 VTPVSSSL-GTQYICNVNHNKPSNTKVKKVPKSCDKTHTCPCPAPALGAPSVFLFP 127  
 QY 342 PKPKDTLMSRTPVETCVVVDVSHEDPEVKFNWYDGVVEVHNATKPREEQNSTYRVVS 401  
 DB 128 PKPKDTLMSRTPVETCVVVDVSHEDPEVKFNWYDGVVEVHNATKPREEQNSTYRVVS 187

QY 402 VLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRREMTKNQVS 461  
 DB 188 VLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVS 247  
 QY 462 LTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFS 521  
 DB 248 LTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFS 307  
 QY 522 CSVMHRLAHNHYTKQSLSLSPGK 544  
 DB 308 CSVMHRLAHNHYTKQSLSLSPGK 330

## RESULT 4

Ig gamma-1 chain C region - synthetic

C:Species: synthetic

C:Date: 06-Jan-1995 #sequence\_revision 17-Mar-1997 #text\_change 19-May-2000

C:Accession: S31866

R:Filpula, D.

submitted to the EMBL Data Library, February 1993

A:Description: Screening method for protein-protein interactions of cloned gene product

A:Reference number: S31866

A:Accession: S31866

A:Molecule type: mRNA

A:Residues: 1-255 &lt;FIL&gt;

A:Cross-references: EMBL:X70421; NID:G33068; PIDN:CAA49866.1; PID:G33069

C:Keywords: immunoglobulin

F:1-22/Region: Escherichia coli outer membrane protein A precursor

F:23-255/Region: human Ig gamma-1 chain C region

Query Match 41.2%; Score 1197; DB 4; Length 255;

Best Local Similarity 96.9%; Pred. No. 8.4e-75;

Matches 219; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 319 RPHCTPCPAPALGAPSVFLFPKPKDTLMSRTPVETCVVVDVSHEDPEVKFNWYVDG 378

DB 30 KTHCTPCPAPALGAPSVFLFPKPKDTLMSRTPVETCVVVDVSHEDPEVKFNWYVDG 89

QY 379 VEVHNATKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 438

DB 90 VEVHNATKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 149

QY 439 QPREPQVYTLPPSRREMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSD 498

DB 150 QPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSD 209

QY 499 GSFFLYSKLTVDKSRWQQGNVFSQVMHEALHNHYTKQSLSLSPGK 544

DB 210 GSFFLYSKLTVDKSRWQQGNVFSQVMHEALHNHYTKQSLSLSPGK 255

RESULT 5

Ig gamma chain C region - chimpanzee

C:Species: Pan troglodytes (chimpanzee)

C:Date: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 16-Jul-1999

C:Accession: PT0207

R:Shrlich, P.H.; Moustafa, Z.A.; Oestberg, L.

Mol. Immunol. 28, 319-322, 1991

A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.

A:Reference number: PT0207; MUID:91287716; PMID:2062315

A:Accession: PT0207

A:Molecule type: mRNA

A:Residues: 1-234 &lt;EHR&gt;

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:48-117/Domain: immunoglobulin homology &lt;IMW&gt;

Query Match 39.6%; Score 1150; DB 2; Length 234;

Best Local Similarity 92.2%; Pred. No. 1.2e-71;

Matches 213; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

QY 313 KPTKARP-----HTCPPCPAEEALCAPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHE 366  
 DB 4 KVDKVEPKSCDTHCTCPCAAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHE 63  
 QY 367 DPEVKFNWYVDGVEVHNAKTKPREEQNTSYRSTVSVLTTLVHQQDLNKGKEYCKKVSNAKLP 426  
 DB 64 DPEVKFNWYVDGVEVHNAKTKPREEQNTSYRSTVSVLTTLVHQQDLNKGKEYCKKVSNAKLP 123  
 QY 427 VPIETKISKAKGQPREPQVITLPPSRDEMTNQVSLTCLVKGYFIPSDIAVEWESNQGPEN 486  
 DB 124 APIETKISKAKGQPREPQVITLPPSRDELTKNQVSLTCLVKGYFIPSDIAVEWESSQGPEN 183  
 QY 487 NYKTPRPVLDSDGSFFLYSKLTVDKSRWQQGNFVSCSVWHEALHNHYTOKS 537  
 DB 184 NYKTPRPVLDSDGSFFLYSKLTVDKSRWQQGNFVSCSVWHEALHNHYTOKS 234

RESULT 6  
 G2HU  
 IG gamma-2 chain C region - human  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Apr-1981 #sequence\_revision 13-Jun-1983 #text\_change 21-Jul-2000  
 C:Accession: A93906; A92809; A90752; A93132; A02148  
 R:Ellison, J.; Hood, L.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982  
 A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con  
 A:Reference number: A93906; MUID:82197621; PMID:6804948  
 A:Accession: A93906  
 A:Molecule type: DNA  
 A:Residues: 1-326 <ELL>  
 A:Cross-references: GB:V00554; GB:J00230; NID:G32759; PIDN:CAB58438.1; PID:G6066056  
 A:Note: Lys-326 is probably removed posttranslationally  
 R:Wang, A.C.; Tung, E.; Fudenberg, H.H.  
 J. Immunol. 125, 1048-1054, 1980  
 A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and f  
 A:Reference number: A92809; MUID:81007873; PMID:6774012  
 A:Contents: myeloma protein Til  
 A:Accession: A92809  
 A:Molecule type: protein  
 A:Residues: 1-19, 'Q', 21-57, 'Z', 59, 'A', 61-193, 'D', 195-325 <WAN>  
 A:Note: Trp-156 is at or near the complement-binding site  
 R:Connell, G.E.; Parr, D.M.; Hofmann, T.  
 Can. J. Biochem. 57, 758-767, 1979  
 A:Title: The amino acid sequences of the three heavy chain constant region domains of a  
 A:Reference number: A90752; MUID:80001357; PMID:113060  
 A:Contents: myeloma protein Zie  
 A:Accession: A90752  
 A:Molecule type: protein  
 A:Residues: 1-24, 'E', 26-57, 'EV', 60-85, 132-171, 'ZZZ', 175, 'B', 177-193, 'D', 195-196, 'Q', 198-  
 A:Note: this sequence has since been revised  
 R:Hofmann, T.; Parr, D.M.  
 Mol. Immunol. 16, 923-925, 1979  
 A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin G  
 A:Reference number: A93132; MUID:80114419; PMID:118920  
 A:Contents: Zie  
 A:Accession: A93132  
 A:Molecule type: protein  
 A:Residues: 238-275 <HOF>  
 R:Hofmann, T.; Parr, D.M.  
 submitted to the Atlas, March 1980  
 A:Reference number: A94591  
 A:Contents: annotation: Zie, revisions to residues 25, 59, 60, and 264-268  
 A:Note: the revised sequence differs from that shown in having 60-Ala and in the amidati  
 ned  
 R:Milstein, C.; Frangione, B.  
 Biochem. J. 121, 217-225, 1971  
 A:Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.  
 A:Reference number: A90253; MUID:72033500; PMID:4940472  
 A:Contents: annotation: myeloma protein Sa, disulfide bonds  
 R:Frangione, B.; Milstein, C.; Pink, J.R.L.  
 Nature 221, 145-148, 1969  
 A:Title: Structural studies of immunoglobulin G.

A:Reference number: A93157; MUID:69064124; PMID:5782707  
 A:Contents: annotation; Sa, disulfide bonds  
 C:Genetics:  
 A:Gene: GDB:IGHG2  
 A:Cross-references: GDB:119338; OMIM:147110  
 A:Map position: 14q32.33-14q32.33  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (/)  
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IMI>  
 F:133-202/Domain: immunoglobulin homology <IM2>  
 F:239-306/Domain: immunoglobulin homology <IM3>  
 F:14/Disulfide bonds: interchain (to light chain) #status experimental  
 F:27-83/140-200,246-304/Disulfide bonds: #status experimental  
 F:102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental  
 F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 39.6%; Score 1149.5; DB 1; Length 326;  
 Best Local Similarity 70.5%; Pred. No. 2.1e-71;  
 Matches 227; Conservative 20; Mismatches 34; Indels 41; Gaps 6;  
 QY 253 KAMTSNVAQDNDKRP-----VYKYP-----GKGQPTL----- 283  
 DB 16 RSTSESTAALGCLVKDYFPPVTVSMNSGALTSVHTFPAVLQSSGLYSLSSVTVFSSN 75  
 QY 284 -GDEGDTLDLYYDEETEGDKVRAITVVKFKAPKPHCTCPCPAPEALGAPSVFLPPP 342  
 DB 76 FGQTQYTCNVDPKPS-SNTKVDKTVKRCCKVE-----CPPCPAPPVAG-PSVFLPPP 124  
 QY 343 KPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQNTSYRVSU 402  
 DB 125 KPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQNTSYRVSU 184  
 QY 403 LTVLHQDLNKGKEYCKKVSNAKLPVPIETKISKAKGQPREPQVITLPPSRDEMTKNQVSL 462  
 DB 185 LTVLHQDLNKGKEYCKKVSNAKLPVPIETKISKAKGQPREPQVITLPPSRDEMTKNQVSL 244  
 QY 463 TCLKVGFYPSDIAVEWESNQGPENNYKTTPVLDSDGSFFLYSKLTVDKSRWQQGNFVSC 522  
 DB 245 TCLKVGFYPSDIAVEWESNQGPENNYKTTPVLDSDGSFFLYSKLTVDKSRWQQGNFVSC 304  
 QY 523 SVMHEALHNHYTOKSLSLSPGK 544  
 DB 305 SVMHEALHNHYTOKSLSLSPGK 326

RESULT 7  
 A23511  
 IG gamma-3 chain C region (allotype G3m(b)) - human  
 C:Species: Homo sapiens (man)  
 C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 23-Jul-1999  
 C:Accession: A23511  
 R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.  
 Nucleic Acids Res. 14, 1779-1789, 1986  
 A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene:  
 A:Reference number: A23511; MUID:86148507; PMID:3081877  
 A:Accession: A23511  
 A:Molecule type: DNA  
 A:Residues: 1-377 <HUC>  
 A:Cross-references: GB:X03604; GB:M12958; NID:G33070; PIDN:CAA27268.1; PID:G577056  
 C:Genetics:  
 A:Gene: GDB:IGHG3  
 A:Cross-references: GDB:119339; OMIM:147120  
 A:Map position: 14q32.33-14q32.33  
 A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 39.3%; Score 1141.5; DB 2; Length 377;  
 Best Local Similarity 63.6%; Pred. No. 9.1e-71;  
 Matches 232; Conservative 19; Mismatches 57; Indels 57; Gaps 8;





Db 327 K 327

## RESULT 10

## G3H0W1

Ig gamma-3 heavy chain disease proteins - human  
 C:Species: Homo sapiens (man)  
 C:Date: 31-Dec-1979 #sequence\_revision 23-Oct-1981 #text\_change 16-Jul-1999  
 C:Accession: A90442; A92219; A90198; A93915; A02149  
 R:Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.  
 Biochemistry 19, 4304-4308, 1980  
 A:Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy-chain disease protein Wis  
 A:Reference number: A90442; MUID:81021548; PMID:6774747  
 A:Contents: heavy chain disease protein Wis  
 A:Accession: A90442  
 A:Molecule type: protein  
 A:Residues: 1-289 <FRA>  
 A:Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interchain A:Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 C:Note: the sequence of residues 42-76 was taken from the reference that follows  
 R:Michaelson, T.E.; Frangione, B.; Franklin, E.C.  
 J. Biol. Chem. 252, 883-889, 1977  
 A:Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication A:Reference number: A92219; MUID:77118561; PMID:402363  
 A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein W A:Accession: A92219  
 A:Molecule type: protein  
 A:Residues: 12-97 <MIC>  
 A:Note: the hinge region in gamma-3 chains is about four times as long as in other gamma-3 chains (12-28)  
 R:Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.  
 Biochem. Biophys. Res. Commun. 71, 907-914, 1976  
 A:Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of the A:Reference number: A90198; MUID:77021516; PMID:823945  
 A:Contents: heavy chain disease protein Zuc, partial sequence corresponding to residues A:Accession: A90198  
 A:Molecule type: protein  
 A:Residues: 59-125, 'EB', 128-226, 228-289 <WOL>  
 A:Note: this protein lacks most of the V region, all of the CH1 region, and part of the A:Alexandri, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood, L. Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982  
 A:Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion A:Reference number: A93915; MUID:82247835; PMID:6808505  
 A:Contents: heavy chain disease protein Omm  
 A:Accession: A93915  
 A:Molecule type: mRNA  
 A:Residues: 12-70; 72-114; 116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-157 A:Note: a carboxyl-terminal Lys is removed posttranslationally  
 A:Note: this sequence may represent an allelic form or another gamma chain subclass C:Comment: The heavy chain disease protein Wis is shown.  
 C:Genetics:  
 A:Gene: GDB:10HG3  
 A:Cross-references: GDB:119339; OMIM:147120  
 A:Map position: 14q32.33-14q32.33  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 F:203-270/Domain: immunoglobulin homology <IMM>  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
 F:6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 38.4%; Score 1115; DB 1; Length 289;  
 Best Local Similarity 90.6%; Pred. No. 4-le-69;  
 Matches 203; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

Qy 320 PHTCPCPAPALGARSVFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGV 379  
 Db 66 PPPCPAPALGARSVFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGV 125

Qy 380 EVNNAKTPREEOYNSTYRVVSVLTVLHQDMLNGKEYCKVSNKALPVIETKISKAKQ 439  
 Db 126 QVHNAKTPREEOYNSTYRVVSVLTVLHQDMLNGKEYCKVSNKALPVIETKISKAKQ 185

Qy 440 PREPOVYTLPPSRBEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLDSDG 499  
 Db 186 PREPOVYTLPPSRBEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLDSDG 245  
 Qy 500 SFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPG 543  
 Db 246 SFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPG 289

## RESULT 11

## I47160

Ig gamma 2b chain constant region - pig (fragment)  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
 C:Accession: I47160  
 R:Kackovics, I.; Sun, J.; Butler, J.E.  
 J. Immunol. 153, 3565-3573, 1994  
 A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of A:Reference number: I47158; MUID:95015845; PMID:7930579  
 A:Accession: I47160  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-328 <KAC>  
 A:Cross-references: EMBL:U03780; NID:9433125; PIDN:AAA52218.1; PID:9433126  
 C:Genetics:  
 A:Gene: Igg2b  
 A:Superfamily: immunoglobulin C region; immunoglobulin homology  
 F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 31.6%; Score 918; DB 2; Length 328;  
 Best Local Similarity 67.5%; Pred. No. 1.6e-55;  
 Matches 170; Conservative 34; Mismatches 36; Indels 12; Gaps 4;

Qy 295 YPEDTEGDKVTRATRVVFPFKAPHTCPCPAPALGAPSVLPFPKPKDTLMISRT 354  
 Db 87 HPATTKVKD-----RVGTKTP-PCPICPACESPG-PSVFIFFPKPKDTLMISRT 136

Qy 355 EVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDMLNGK 414  
 Db 137 QVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDMLNGK 196

Qy 415 EYKCKVSNKALPVIETKISKAKQPREPPQPVYTLPPSREEMTKNQVSLTCLVKGFYPSDI 474  
 Db 197 EFKCKVSNKALPVIETKISKAKQPREPPQPVYTLPPSREEMTKNQVSLTCLVKGFYPSDI 256

Qy 475 AVEWESNGQ--PENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNH 532  
 Db 257 DVEMQRNQPPEEGNYRTTTPQDQVDGTYFLYSKFSVDKASWQGGGIFQCAVMEALHNH 316

Qy 533 YTKSLSLSPGK 544  
 Db 317 YTKSLSLSPGK 328

## RESULT 12

## I47159

Ig gamma 2a chain constant region - pig (fragment)  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
 C:Accession: I47159  
 R:Kackovics, I.; Sun, J.; Butler, J.E.  
 J. Immunol. 153, 3565-3573, 1994  
 A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of A:Reference number: I47158; MUID:95015845; PMID:7930579  
 A:Accession: I47159  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-328 <KAC>  
 A:Cross-references: EMBL:U03779; NID:9433123; PIDN:AAA52217.1; PID:9433124  
 C:Genetics:  
 A:Gene: Igg2a  
 A:Superfamily: immunoglobulin C region; immunoglobulin homology  
 F:133-202/Domain: immunoglobulin homology <IMM>

C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin

F:20-82/Domain: immunoglobulin homology <IM1>  
F:130-199/Domain: immunoglobulin homology <IM2>  
F:236-303/Domain: immunoglobulin homology <IM3>  
F:173/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 31.4%; Score 913; DB 1; Length 323;  
Best Local Similarity 66.3%; Pred. NO. 3.4e-55;  
Matches 167; Conservative 31; Mismatches 40; Indels 14; Gaps 2;

QY 295 YPEETEGDKVRATRVVVFPTKARPHTC--PPCPAPALGAPSVFLFPPPKDPTLMISR 352  
DB 84 HPATNIKVDKTV-----PSTCSKPTCPPELLGGPSVFLFPPPKDPTLMISR 131  
QY 353 TPEVTCVVVDVSHEDPRVKFNWYDGVENVNAKTKPREEQYNSTYRVVSVLTVHLQDWLN 412  
DB 132 TPEVTCVVVDVSQDDPEVQFTWYINNEQVTRAPPLREQQFNSTIRVVSUTLPITHQDWLR 191  
QY 413 GKEYCKVSNKALPVPILKTIISKAKGQPRQVTVLPPSREMTKNQVSLTCLVKGYGYP 472  
DB 192 GKEFKCKVHKALPAPIEKTIISKAKGQPLEPKVYTMGPPREELSSRSVSLTCMNGEYPS 251  
QY 473 DIAVWESNGQPNENYKTTTPVLSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNH 532  
DB 252 DISVBEWKNKGKAEADNKITTPAVLSDGSFFLYNKLVSPTSEWQRGDVFTCSVMHEALHNH 311  
QY 533 YTQKSLSLSPGK 544  
DB 312 YTQKSLSLSPGK 323

RESULT 14  
I47162  
IG Gamma 4 chain constant region - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
C:Accession: I47162  
R:Kaczkovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IGG identified from the cDNA sequences of  
A:Reference number: I47158; MUID:95015845; PMID:7930579  
A:Accession: I47162  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-277 <KAC>  
A:Cross-references: EMBL:U03782; NID:G433129; PIDN:AAA52220.1; PID:G433130  
C:Genetics:  
A:Gene: IG4  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:82-151/Domain: immunoglobulin homology <IM>

Query Match 31.3%; Score 911; DB 2; Length 277;  
Best Local Similarity 67.1%; Pred. NO. 3.7e-55;  
Matches 169; Conservative 33; Mismatches 38; Indels 12; Gaps 4;

QY 295 YPEETEGDKVRATRVVVFPTKARPHTCPPCPAPALGAPSVFLFPPPKDPTLMISRT 354  
DB 36 HPATNIKVDK-----RVGKTKP-PCPICPACGPG-PSAIFPPPKDPTLMISRT 85  
QY 355 EYTCVVVDVSHEDPRVKFNWYDGVENVNAKTKPREEQYNSTYRVVSVLTVHLQDWLNK 414  
DB 86 KYTCVVVDVSQENPEVQFSWYDGVENVTAQTRPEKQFNSTYRVWSYLPVQHODWLNK 145  
QY 415 EYCKVSNKALPVPILKTIISKAKGQPRQVTVLPPSREMTKNQVSLTCLVKGYPSDI 474  
DB 146 EFKCKVNNKDPAPITRIISKAKGQTRPEQVYTLUPPTPELSRSKVTLCVTGYFPDI 205  
QY 475 AVYEWESNGQ--PENNYKTTTPVLSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNH 532  
DB 206 DVEWQNGQPEPEGNYRTTPQDDVDGTYFFLYSKLAVDVKASWQRGDTFQCAVMHEALHNH 265  
QY 533 YTQKSLSLSPGK 544  
DB 312 YTQKSLSLSPGK 323



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:16:09 ; Search time 15.1814 Seconds  
(without alignments)  
1685.124 Million cell updates/sec

Title: US-10-068-426-1  
Perfect score: 2906  
Sequence: 1 MPLLLLLLSPLPHPIC.....MHEALHHYTKSLSPGK 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1693	58.3	626	1 GPBA HUMAN
2	1197.5	41.2	330	1 GCI_HUMAN
3	1149.5	39.6	326	1 GC2_HUMAN
4	1137	39.1	327	1 GC4_HUMAN
5	1120	38.5	290	1 GC1_HUMAN
6	913	31.4	323	1 GC_RABIT
7	887	30.5	329	1 GC2_CAVPO
8	843	29.0	329	1 GC3_MOUSE
9	832	28.6	398	1 GC3M_MOUSE
10	831	28.6	326	1 GCI_RAT
11	828	28.5	333	1 GCB_RAT
12	823	28.3	329	1 GCC_RAT
13	820.5	28.2	324	1 GCI_MOUSE
14	815.5	28.1	393	1 GC1M_MOUSE
15	812	27.9	322	1 GCA_RAT
16	805	27.7	330	1 GCAA_MOUSE
17	802	27.6	335	1 GCAB_MOUSE
18	800	27.5	399	1 GCAM_MOUSE
19	767.5	26.4	336	1 GCB_MOUSE
20	762.5	26.2	405	1 GCBM_MOUSE
21	359.5	12.4	421	1 EPC_MOUSE
22	358	12.3	454	1 MUC_HUMAN
23	356	12.3	391	1 MUC2_HUMAN
24	354	12.2	429	1 EPC_RAT
25	353.5	12.2	428	1 EPC_MOUSE
26	352	12.1	455	1 MUC_MOUSE
27	342	11.8	476	1 MUCM_MOUSE
28	338	11.6	458	1 MUC_RABIT
29	331.5	11.4	450	1 MUC_CANFA
30	329	11.3	454	1 MUC_MESAU
31	328	11.3	479	1 MUCM_RABIT
32	322	11.1	457	1 MUC_SUNMU
33	301	10.4	347	1 A2GL_HUMAN

34	295	10.2	438	1	HVC2_HETFR	P23085 heterodontu
35	293	10.1	299	1	ALC_RABIT	P01879 erylcolagus
36	289.5	10.0	446	1	MUC_CHICK	P01875 gallus gall
37	286	9.8	353	1	ALCI_HUMAN	P01876 homo sapien
38	283	9.7	438	1	HVC3_HETFR	P23087 heterodontu
39	279	9.6	353	1	ALCI_GORGO	P20758 gorilla gor
40	277.5	9.5	340	1	ALC2_HUMAN	P01877 homo sapien
41	274	9.4	370	1	HVC1_HETFR	P23084 heterodontu
42	274	9.4	461	1	HVC3_HETFR	P23088 heterodontu
43	271	9.3	393	1	HVC3_HETFR	P23086 heterodontu
44	264.5	9.1	646	1	FLRI_HUMAN	Q9nzu1 homo sapien
45	259	8.9	344	1	ALC_MOUSE	P01878 mus musculu

## ALIGNMENTS

RESULT 1  
GPBA\_HUMAN  
ID GPBA\_HUMAN STANDARD; PRT; 626 AA.  
AC P07359; O9HDC7;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1988 (Rel. 07, last sequence update)  
DE Platelet glycoprotein Ib alpha chain precursor (Glycoprotein Ibalpha)  
DE (GP-Ib alpha) (GP1BA) (CD42B-alpha) (CD42B) [Contains:  
DE Glycocalicin].  
GN GP1BA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=87289655; PubMed=3303030;  
RA Lopez J.A., Chung D.W., Fujikawa K., Hagen F.S., Papayannopoulou T.,  
RA Roth G.J.;  
RT "Cloning of the alpha chain of human platelet glycoprotein Ib: a  
RT transmembrane protein with homology to leucine-rich alpha 2-  
RT glycoprotein.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:5615-5619(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89025874; PubMed=2845978;  
RA Wenger R.H., Kieffer N., Wicki A.N., Clemetson K.J.;  
RT "Structure of the human blood platelet membrane glycoprotein Ib alpha  
RT gene.";  
RL Biochem. Biophys. Res. Commun. 156:389-395(1988).  
RN [3]  
RP SEQUENCE FROM N.A., AND VARIANT PHE-86.  
RX MEDLINE=22034223; PubMed=12038791;  
RA Matsubara Y., Murata M., Moriki T., Yokoyama K., Watanabe N.,  
RA Nakajima H., Handa M., Kawano K., Aoki N., Yoshino H., Ikeda Y.;  
RT "A novel polymorphism, 70Leu/Phe, disrupts a consensus Leu residue  
RT within the leucine-rich repeat sequence of platelet glycoprotein  
RT Ibalpha.";  
RL Thromb. Haemost. 87:867-872(2002).  
RN [4]  
RP SEQUENCE FROM N.A., AND VARIANT MET-161.  
RX Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Yi Q.,  
RA Nickerson D.A.;  
RN Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE OF 17-315.  
RX MEDLINE=87289654; PubMed=3497398;  
RA Titani K., Takio K., Handa M., Ruggeri Z.M.;  
RT "Amino acid sequence of the von Willebrand factor-binding domain of  
RT platelet membrane glycoprotein Ib.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:5610-5614(1987).  
RN [6]  
RP DISULFIDE BONDS.  
RX MEDLINE=91301149; PubMed=2070794;  
RA Hess D., Schaller J., Rickli E.E., Clemetson K.J.;

RT "Identification of the disulphide bonds in human platelet  
 RT glycofibrin.";  
 RL Eur. J. Biochem. 199;389-393(1991).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).  
 RA MEDLINE=92171415; PubMed=12183630;  
 RX Huizinga E.G., Tsuji S., Romijn R.A., Schiphorst M.E., de Groot P.G.,  
 RA Sixma J.J., Gros P.;  
 RT "Structures of glycoprotein Ib alpha and its complex with von  
 RT Willebrand factor A1 domain.";  
 RL Science 297:1176-1179(2002).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 17-304, AND SULFATION OF  
 RP TYR-292; TYR-294 AND TYR-295.  
 RA MEDLINE=92217968; PubMed=12087105;  
 RX Uff S., Clemetson J.M., Harrison T., Clemetson K.J., Ensley J.;  
 RT "Crystal structure of the platelet glycoprotein Ib alpha N-terminal  
 RT domain reveals an unmasking mechanism for receptor activation.";  
 RL J. Biol. Chem. 277:35657-35663(2002).  
 RN [9]  
 RP VARIANT SIBA MET-161.  
 RA MEDLINE=92265982; PubMed=1586750;  
 RX Murata M., Furihata K., Ishida F., Russell S.R., Ware J.,  
 RA Ruggeri Z.M.;  
 RT "Genetic and structural characterization of an amino acid dimorphism  
 RT in glycoprotein Ib alpha involved in platelet transfusion  
 RT refractoriness.";  
 RL Blood 79:3086-3090(1992).  
 RN [10]  
 RP VARIANT BSS PHE-73.  
 RA MEDLINE=92110577; PubMed=1730088;  
 RX Miller J.L., Lyle V.A., Cunningham D.;  
 RT "Mutation of leucine-57 to phenylalanine in a platelet glycoprotein  
 RT Ib alpha leucine tandem repeat occurring in patients with an  
 RT autosomal dominant variant of Bernard-Soulier disease.";  
 RL Blood 79:439-446(1992).  
 RN [11]  
 RP VARIANT BSS VAL-172.  
 RA MEDLINE=93388851; PubMed=7690774;  
 RX Ware J., Russell S.R., Marchese P., Murata M., Mazzucato M.,  
 RA de Marco L., Ruggeri Z.M.;  
 RT "Point mutation in a leucine-rich repeat of platelet glycoprotein Ib  
 RT alpha resulting in the Bernard-Soulier syndrome.";  
 RL J. Clin. Invest. 92:1213-1220(1993).  
 RN [12]  
 RP VARIANT BSS SER-225.  
 RA MEDLINE=95118882; PubMed=7819107;  
 RX Simsek S., Norris P., Lozano M., Pico M., von Dem Borne A.E.G.K.,  
 RA Ribera A., Gallardo D.;  
 RT "Cys209 Ser mutation in the platelet membrane glycoprotein Ib alpha  
 RT gene is associated with Bernard-Soulier syndrome.";  
 RL Br. J. Haematol. 88:839-844(1994).  
 RN [13]  
 RP VARIANT PSEUDO-VWD VAL-249.  
 RA MEDLINE=91271273; PubMed=2052556;  
 RX Miller J.L., Cunningham D., Lyle V.A., Finch C.N.;  
 RT "Mutation in the gene encoding the alpha chain of platelet  
 RT glycoprotein Ib in platelet-type von Willebrand disease.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:4761-4765(1991).  
 RN [14]  
 RP VARIANT PSEUDO-VWD VAL-249.  
 RA MEDLINE=93253059; PubMed=8486780;  
 RX Murata M., Russell S.R., Ruggeri Z.M., Ware J.;  
 RT "Expression of the phenotypic abnormality of platelet-type von  
 RT Willebrand disease in a recombinant glycoprotein Ib alpha fragment.";  
 RL J. Clin. Invest. 91:2133-2137(1993).  
 RN [15]  
 RP VARIANT PSEUDO-VWD VAL-255.  
 RA MEDLINE=93214031; PubMed=8384898;  
 RX Russell S.D., Roth G.J.;  
 RT "Pseudo-von Willebrand disease: a mutation in the platelet  
 RT glycoprotein Ib alpha gene associated with a hyperactive surface  
 RT receptor.";  
 RL Blood 81:1787-1791(1993).  
 RN [16]  
 RP VARIANT BSS LEU-195 DEL.  
 RA MEDLINE=95178321; PubMed=7873390;  
 RX de la Salle C., Baas M.-J., Lanza F., Schwartz A., Hanau D.,  
 RA Chevalier J., Gachet C., Briquel M.-E., Cazenave J.-P.;  
 RT "A three-base deletion removing a leucine residue in a leucine-rich  
 RT repeat of platelet glycoprotein Ib alpha associated with a variant of  
 RT Bernard-Soulier syndrome (Nancy I).";  
 RL Br. J. Haematol. 89:386-396(1995).  
 RN [17]  
 RP VARIANT BSS ARG-81.  
 RA MEDLINE=98303759; PubMed=9639514;  
 RX Kenny D., Jonsson O.G., Mordecki P.A., Montgomery R.R.;  
 RT "Naturally occurring mutations in glycoprotein Ib alpha that result in  
 RT defective ligand binding and synthesis of a truncated protein.";  
 RL Blood 92:175-183(1998).  
 RN [18]  
 RP VARIANTS HIS-72 AND MET-161.  
 RA MEDLINE=99318093; PubMed=10391209;  
 RX Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.;  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 RT of human genes.";  
 RL Nat. Genet. 22:231-238(1999).  
 RN [19]  
 RP ERRATUM.  
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.;  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 RT of human genes.";  
 RL Nat. Genet. 22:231-238(1999).  
 RN [20]  
 RP VARIANT BSS PRO-145.  
 RA MEDLINE=99189763; PubMed=10089893;  
 RX Koskela S., Partanen J., Salmi T.T., Kekomaki R.;  
 RT "Molecular characterization of two mutations in platelet glycoprotein  
 RT (GP) Ib alpha in two Finnish Bernard-Soulier syndrome families.";  
 RL Eur. J. Haematol. 62:160-168(1999).  
 CC -!- FUNCTION: GP-Ib, a surface membrane protein of platelets,  
 CC participates in the formation of platelet plugs by binding to the  
 CC A1 domain of von Willebrand factor, which is already bound to the  
 CC subendothelium.  
 CC -!- SUBUNIT: Heterodimer composed of GP-Ib alpha and beta; disulfide  
 CC linked. GP-IX is complexed with the GP-Ib heterodimer via a non  
 CC covalent linkage.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- PTM: Glycosylation, which is approximately coextensive with the  
 CC extracellular part of the molecule, is cleaved off by calpain  
 CC during platelet lysis.  
 CC -!- POLYMORPHISM: Position 161 is associated with platelet-specific  
 CC allotypic Siba. Siba(-) has Thr-161 and Siba(+) has Met-161.  
 CC Siba is involved in neonatal alloimmune thrombocytopenia (NATP).  
 CC -!- DISEASE: Defects in GPIBA are one of the causes of Bernard-Soulier  
 CC syndrome (BSS). BSS patients have unusually large platelets and  
 CC have a clinical bleeding tendency.  
 CC -!- DISEASE: Defects in GPIBA are one of the causes of von Willebrand  
 CC disease (VWD) known as platelet-type von Willebrand disease or  
 CC pseudo-von Willebrand disease. This autosomal dominant bleeding  
 CC disorder is caused by an increased affinity of GP-Ib for soluble  
 CC vWF resulting in impaired hemostatic function due to the removal  
 CC of vWF from the circulation.  
 CC -!- MISCELLANEOUS: Platelet activation apparently involves disruption  
 CC of the macromolecular complex of GP-Ib with the platelet  
 CC glycoprotein IX (GP-IX) and dissociation of GP-Ib from the actin-  
 CC binding protein.  
 CC -!- MISCELLANEOUS: Binding sites for von Willebrand factor and  
 CC thrombin (the latter site with unknown function) are in the amino-  
 CC terminal part of the molecule.  
 CC -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.  
 CC -----





DR MM; 147100; --  
 DR GO; GO:0005624; C:membrane fraction; NAS.  
 DR GO; GO:0003623; F:antigen binding activity; TAS.  
 DR GO; GO:0006955; P:immune response; NAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig\_c1.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00407; Igc1; 2.  
 DR PROSITE; PS00835; IG LIKE; 3.  
 DR PROSITE; PS00290; IG\_MHC; 2.  
 DR KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 3D-structure.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 98 CH1.  
 FT DOMAIN 99 110 HINGE.  
 FT DOMAIN 111 223 CH2.  
 FT DOMAIN 224 330 CH3.  
 FT DISULFID 27 83  
 FT DISULFID 103 103 INTERCHAIN (WITH LIGHT CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN).  
 FT DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).  
 FT DISULFID 144 204  
 FT DISULFID 250 308  
 FT CARBOHYD 180 180  
 FT MOD\_RES 330 330  
 FT VARIANT 97 97  
 FT VARIANT 239 239  
 FT VARIANT 241 241  
 FT STRAND 122 126  
 FT HELIX 130 134  
 FT TURN 136 137  
 FT STRAND 141 147  
 FT STRAND 157 162  
 FT TURN 163 164  
 FT STRAND 165 166  
 FT TURN 168 171  
 FT STRAND 176 179  
 FT TURN 180 181  
 FT STRAND 182 190  
 FT HELIX 193 197  
 FT TURN 198 199  
 FT STRAND 202 207  
 FT TURN 209 210  
 FT STRAND 215 219  
 FT STRAND 227 227  
 FT STRAND 230 234  
 FT STRAND 238 242  
 FT STRAND 245 256  
 FT STRAND 260 265  
 FT STRAND 270 270  
 FT STRAND 274 276  
 FT STRAND 280 281  
 FT TURN 283 284  
 FT STRAND 287 296  
 FT STRAND 297 301  
 FT TURN 302 303  
 FT STRAND 305 312  
 FT TURN 313 314  
 FT TURN 316 317  
 FT STRAND 320 325  
 SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;  
 Query Match 41.2%; Score 1197.5; DB i; Length 330;  
 Best Local Similarity 73.1%; Pred. No. 4.8e-78;  
 Matches 236; Conservative 11; Mismatches 37; Indels 39; Gaps 6;  
 QY 253 KAMTSNVAIVCDNSDKFP-----VYKYPGKGCPTLGDGCTDLYDYYPE 297  
 DB 16 KSTSGGTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPA-----VLOSSG---LYSLSSV 68

QY 298 EDTGDKVRAIRTVV-----KPPTKARP-----HTCPPCAPALGAPSVFLFP 341  
 DB 69 VTFPSSL-CTQTYICNVNHPKSPNTKVDKKVBPKSCDKTHTCPCPCAPALGAPSVFLFP 127  
 QY 342 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVS 401  
 DB 128 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVS 187  
 QY 402 VLTVLHQMGLNGKEYCKKSNKALPVPVIEKTIISKAKQGPPEQVYTLPPSREMTKNQVS 461  
 DB 188 VLTVLHQMGLNGKEYCKKSNKALPVPVIEKTIISKAKQGPPEQVYTLPPSREMTKNQVS 247  
 QY 462 LTCLVKGFPYSDIAVWESNGQPNKYKTPPVLDSDGSEFLYSKLTVDKSRWQCGNVFS 521  
 DB 248 LTCLVKGFPYSDIAVWESNGQPNKYKTPPVLDSDGSEFLYSKLTVDKSRWQCGNVFS 307  
 QY 522 CSVMHEALHNHYTQKSLSLSPGK 544  
 DB 308 CSVMHEALHNHYTQKSLSLSPGK 330  
 RESULT 3  
 GC2\_HUMAN  
 ID GC2\_HUMAN STANDARD; PRT; 326 AA.  
 AC P01859;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE IG gamma-2 chain C region.  
 GN IGHG2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE OF 2-326 FROM N.A.  
 RX MEDLINE=82197621; PubMed=6804948;  
 RA Ellison J.W., Hood L.E.;  
 RT "Linkage and sequence homology of two human immunoglobulin gamma  
 heavy chain constant region genes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).  
 RN [2]  
 RP SEQUENCE OF 88-115 FROM N.A.  
 RC TISSUE=Fetal liver;  
 RX MEDLINE=83001943; PubMed=6811139;  
 RA Takahashi N., Ueda S., Obata M., Nikaido T., Nakai S., Honjo T.;  
 RT "Structure of human immunoglobulin gamma genes: implications for  
 evolution of a gene family.";  
 RL Cell 29:671-679(1982).  
 RN [3]  
 RP SEQUENCE OF 99-177 AND 310-326 FROM N.A.  
 RC TISSUE=Fetal liver;  
 RX MEDLINE=84235992; PubMed=6329676;  
 RA Krawinkel U., Rabbitts T.H.;  
 RT "Comparison of the hinge-coding segments in human immunoglobulin gamma  
 heavy chain genes and the linkage of the gamma 2 and gamma 4 subclass  
 genes.";  
 RL EMBO J. 1:403-407(1982).  
 RN [4]  
 RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).  
 RX MEDLINE=81007873; PubMed=6774012;  
 RA Wang A.-C., Tung E., Fudenberg H.H.;  
 RT "The primary structure of a human IGG2 heavy chain: genetic,  
 evolutionary, and functional implications.";  
 RL J. Immunol. 125:1048-1054(1980).  
 RN [5]  
 RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).  
 RX MEDLINE=80001357; PubMed=113060;  
 RA Connell G.E., Parr D.M., Hofmann T.;  
 RT "The amino acid sequences of the three heavy chain constant region  
 domains of a human IGG2 myeloma protein.";  
 RL Can. J. Biochem. 57:758-767(1979).

RP SEQUENCE OF 238-275 (ZIE).  
RX MEDLINE=80114119; PubMed=118920;  
RA Hofmann T., Parr D.M.;  
RT "A note of the amino acid sequence of residues 381-391 of human  
immunoglobulin gamma chains."  
RL Mol. Immunol. 16:923-925(1979).  
RN [7].  
RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).  
RA Hofmann T., Parr D.M.;  
RL Submitted (MAR-1980) to the PIR data bank.  
RN [8].  
RP SEQUENCE OF 1-121 (DOT).  
RX MEDLINE=95255298; PubMed=7737190;  
RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;  
RT "Characterization of the two unique human anti-flavin monoclonal  
immunoglobulins."  
RL Eur. J. Biochem. 228:886-893(1995).  
RN [9].  
RP DISULFIDE BONDS.  
RX MEDLINE=72033500; PubMed=4940472;  
RA Milstein C., Frangione B.;  
RT "Structural studies of immunoglobulin G."  
RL Nature 221:145-148(1969).  
RN [10].  
RP This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
RN EMBL; J00230; AAB59393.1; -;  
DR PIR; A93906; G2HU;  
DR HSP; P01857; IFC1;  
DR Genew; HGNC:5526; IGHG2.  
DR MIM; 147110; -;  
DR GO; GO:0005624; C:membrane fraction; NAS.  
DR GO; GO:0003823; F:antigen binding activity; TAS.  
DR GO; GO:0006955; P:immune response; NAS.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003597; Ig-cl.  
DR InterPro; IPR003006; Ig\_MHC.  
DR Pfam; PF00047; Ig; 3.  
DR SMART; SM00407; IGH1; 2.  
DR PROSITE; PS50835; IG\_LIKE; 3.  
DR PROSITE; PS00290; IG\_MHC; 2.  
DR Immunoglobulin domain; Immunoglobulin C region.  
FT NON TER 1 1  
FT DOMAIN 1 98 CHI.  
FT DOMAIN 99 110 HINGE.  
FT DOMAIN 111 219 CH2.  
FT DOMAIN 220 326 CH3.  
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 27 83  
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 140 200  
FT DISULFID 246 304  
FT SITE 156 156 AT OR NEAR THE COMPLEMENT-BINDING SITE.  
FT MOD RES 326 326 REMOVED POST-TRANSLATIONALLY (PROBABLE).  
FT VARIANT 60 60 S -> A (IN MYELOMA PROTEINS IIL & ZIE).  
FT CONFLICT 109 109 /FTid=VAR\_003889.  
C -> S (IN REF. 3).

SQ SEQUENCE 326 AA; 35084 MW; 8310878C6878CF9C CRC64;  
Query Match 39.6%; Score 1149.5; DB 1; Length 326;  
Best Local Similarity 70.5%; Pred. No. 1.2e-74;  
Matches 227; Conservative 20; Mismatches 34; Indels 41; Gaps 6;  
QY 253 KAMTSNVASVQCDNSDKFP-----YKYP-----GKGCPYL----- 283  
DB 16 RSTSESTAALGLVKDYFPFVTVSMNSGALTSVHTFPAVLQSSGLYSLSVTVTPSN 75  
QY 284 -GDEGTDLDYYPEDTEGDKVRRTRTVVKKPHHTCCPCPAPBALGAPSVFLPPP 342  
DB 76 FGQTQVTCNVDHKP-SNTKVDKTKVERKCCVE-----CPPCPAPPVAG-PSVFLPPP 124  
QY 343 KPQDTLMISRTPEVTCVAVDVSHEDPEVKFNVDGVEVHNKTKPREBOYNSTYRVSV 402  
DB 125 KPQDTLMISRTPEVTCVAVDVSHEDPEVQFNVDGVEVHNKTKPREBOYNSTYRVSV 184  
QY 403 LTVLHQDWLNGKEYKCKVSNKALPVPKEIKTISKAKGQPREPOVYTLPPSREEMTKNQVSL 462  
DB 185 LTVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTKGQPREPOVYTLPPSREEMTKNQVSL 244  
QY 463 TCVLVKGFYSDIAVEWESNGQPENNYKTTTPVLDSDGSPFLYSKLTVDKSRHQQGVFSC 522  
DB 245 TCVLVKGFYSDIAVEWESNGQPENNYKTTTPMLDSDGSPFLYSKLTVDKSRHQQGVFSC 304  
QY 523 SVMHREALNHYTKSLSLSPGK 544  
DB 305 SVMHREALNHYTKSLSLSPGK 326  
RESULT 4  
GC4\_HUMAN STANDARD; PRT; 327 AA.  
ID\_GC4\_HUMAN  
AC P01861; 1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Ig gamma-4 chain C region.  
GN IGHG4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83157104; PubMed=6299662;  
RA Ellison J.W., Buxbaum J.N., Hood L.E.;  
RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene."  
RL DNA 1:11-18(1981).  
RN [2]  
RP SEQUENCE OF 1-30 AND 81-326.  
RX MEDLINE=70207560; PubMed=4192699;  
RA Pink J.R.L., Burtley S.H., de Vries G.M., Milstein C.;  
RT "Human immunoglobulin subclasses. Partial amino acid sequence of the  
constant region of a gamma 4 chain."  
RL Biochem. J. 117:33-47(1970).  
RN [3]  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
RN EMBL; K01316; AAB59394.1; ALT\_INIT.  
DR PIR; A90933; G4HU.  
DR PDB; 1ADQ; 16-SEP-98.  
DR Genew; HGNC:5528; IGHG4.  
DR MIM; 147130; -;  
DR GO; GO:0005624; C:membrane fraction; NAS.  
DR GO; GO:0003823; F:antigen binding activity; TAS.

```
DR GO:0006955; P:immune response; NAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig_c1.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; IGL1; 2.
DR PROSITE: PS00835; IG_LIKE; 3.
DR PROSITE: PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region; 3D-structure.
FT NON_TER 1
FT DOMAIN 1 98
FT DOMAIN 99 110
FT DOMAIN 111 220
FT DOMAIN 221 327
FT DOMAIN 328 347
FT DISULFID 14 14
FT DISULFID 27 83
FT DISULFID 106 106
FT DISULFID 109 109
FT DISULFID 141 201
FT DISULFID 247 305
SQ SEQUENCE 327 AA; 35940 MW; 3EDBD811EF208E7A CRC64;

Query Match 39.1%; Score 1137; DB 1; Length 327;
Best Local Similarity 75.1%; Pred. No. 9.4e-74;
Matches 226; Conservative 17; Mismatches 42; Indels 16; Gaps 5;

QY 246 WKQGVVDVKAMTSNVAQVC--DNSDKFPVVKYFGKGCFTLGDGDDTDLDYDYPEEDTEGD 303
DB 41 WNSG----ALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTGYTCNVDHKP-SNTKVD 95
QY 304 KVRATRTVVKFPTKARPHCTPCPPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDV 363
DB 96 K-----RVESYKGP-PCPSCAPEFLGSPVFLFPKPKDTLMISRTPEVTCVVVDV 146
QY 364 SHEDPEVKFMYVDGVEVHNAKTPREQKNSTYRVVSVLTVLHQDWLNGKEYKCKVSNK 423
DB 147 SQEDPEVQFMYVDGVEVHNAKTPREQKNSTYRVVSVLTVLHQDWLNGKEYKCKVSNK 206
QY 424 ALPVPKEKTSKAGQPREQVYTPPSRSEMTKNQVSLTCLVKGFYPSDIAVWESNGQ 483
DB 207 GLPSSIEKTSKAGQPREQVYTPPSRSEMTKNQVSLTCLVKGFYPSDIAVWESNGQ 266
QY 484 PENNYKTPPVLDSDGSGFFLYSKLTVDKSRWQQGNFVSCSWMEALHNHYTKSLSPG 543
DB 267 PENNYKTPPVLDSDGSGFFLYSLRTLVDKSRWQQGNFVSCSWMEALHNHYTKSLSLG 326
QY 544 K 544
DB 327 K 327

RESULT 5
GC3_HUMAN
ID GC3_HUMAN STANDARD; PRT; 290 AA.
AC P01860;
DT 21-JUL-1986 (Rel. 01, Created)
DI 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ig gamma-3 chain C region (Heavy chain disease protein) (HDC).
GN IGHG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE (DISEASE PROTEIN WIS).
RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;
RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:
RL gamma 3 heavy-chain disease protein Wis.";
RL Biochemistry 19:4304-4308(1980).
RN [2]
RP REVISIONS TO 12-97 (PROTEIN WIS).

RX MEDLINE=77118561; PubMed=402363;
RA Michaelson T.E., Frangione B., Franklin E.C.;
RT "Primary structure of the 'hinge' region of human IgG3. Probable
RT quadruplication of a 15-amino acid residue basic unit.";
RL J. Biol. Chem. 252:883-889(1977).
RN [3]
RP REVISIONS TO 59-289 (PROTEIN WIS/DISEASE PROTEIN ZUC).
RX MEDLINE=77021516; PubMed=823945;
RA Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;
RT "The amino acid sequence of 'heavy chain disease' protein ZUC.
RT Structure of the Fc fragment of immunoglobulin G3.";
RL Biochem. Biophys. Res. Commun. 71:907-914(1976).
RN [4]
RP SEQUENCE FROM N.A. (DISEASE PROTEIN OMM).
RX MEDLINE=82247835; PubMed=6808505;
RA Alexander A., Steinmetz M., Barritault D., Frangione B.,
RA Franklin E.C., Hood L., Buxbaum J.N.;
RT "Gamma Heavy chain disease in man: cDNA sequence supports partial
RT gene deletion model.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:3260-3264(1982).
CC -!- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA
CC INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11
CC NORMALLY PRESENT IN THE HINGE REGION.
CC -!- MISCELLANEOUS: THE HEAVY CHAIN DISEASE PROTEIN WIS IS SHOWN.
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE
CC REF. 2.
CC -!- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION
CC AND ALL OF THE CH1 REGION.
CC -!- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL
CC OF THE CH1 REGION, AND PART OF THE HINGE COMPARED WITH NORMAL
CC GAMMA-3 HEAVY CHAINS.
CC -!- MISCELLANEOUS: DISEASE PROTEIN OMM MAY REPRESENT AN ALLELIC FORM
CC OR ANOTHER GAMMA CHAIN SUBCLASS.
CC -!- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR
CC TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE
CC IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE
CC SEGMENT (12-28).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/ebis/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; J00231; AAA52805.1; ALT_SEQ.
CC HSP; P01857; IFC1.
CC Genew; HGNC:5527; IGHG3.
CC MIM; 147120; -.
CC GO:0005624; C:membrane fraction; NAS.
CC GO:0003823; F:antigen binding activity; TAS.
CC GO:0006955; F:immune response; NAS.
CC InterPro: IPR007110; Ig-like.
CC InterPro: IPR003597; Ig_c1.
CC Pfam: PF00047; Ig; 2.
CC SMART: SM00407; IGL1; 1.
CC PROSITE: PS00835; IG_LIKE; 2.
CC PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat;
KW Pyridone carboxylic acid.
FT DOMAIN 12 73
FT DOMAIN 74 183
FT DOMAIN 184 289
FT REPEAT 29 43
FT REPEAT 44 58
FT REPEAT 59 73
FT MOD_RES 1 1
FT CARBOHYD 6 6
FT DISULFID 7 7
FT DISULFID 24 24
FT DISULFID 27 27
FT PYRROLIDONE CARBOXYLIC ACID.
FT N-LINKED (GLCNAC...).
FT INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT INTERCHAIN (WITH HEAVY CHAIN DIMER).
```



Db 84 HPAINTKVDKTA-----PSTCKPTCPPELLGGPSVFIFPPKPKDTLMISR 131

Qy 353 TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVWSVLTVLHQDWLN 412

Db 132 TPEVTCVVVDVSDPEVQFTWYINNEQVARTAPPRLREQQFNSTIRVWSTLPITHQDWLR 191

Qy 413 GKEVKCKVSKNALKPVIETKISKAKGQPREPOVYITLPPSEEMTKNQVSLTCLAVKGFPYS 472

Db 192 GKFKCKVHNAKPAPIETKISKARGQPLEPKPVYTMGPPEELSSRSVSLTCMINGFPYS 251

Qy 473 DIAVESNGOPENNYKTTTPVLDSGCSFFLYSKLTVDKSRWQGNVFCVSVNHEALHNH 532

Db 252 DISVEWKGKGAEDNYKTTTPAVLDSGCSFFLYSKLSVPTSEWQSGDVFICSVNHEALHNH 311

Qy 533 YTKSKLSLSPGK 544

Db 312 YTKSKISRSPPGK 323

RESULT 7

GC2\_CAVPO STANDARD; PRT; 329 AA.

ID - GC2\_CAVPO

AC P01862;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig gamma-2 chain C region.

OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.

OX NCBI\_TaxID=10141;

RN [1]

RP SEQUENCE OF 1-3.

RA Trischmann T.M.;

RL Submitted (APR-1975) to the PIR data bank.

RN [2]

RP SEQUENCE OF 4-68.

RX MEDLINE=71058471; PubMed=5538606;

PA Birhrein B.K., Husain O.Z., Cebr J.J.;

RT "Structure of heavy chain from strain 13 guinea pig

RT immunoglobulin-G(2). 3. Amino acid sequence of the region around the

RT half-cystine joining heavy and light chains.";

RN Biochemistry 10:18-25(1971).

RN [3]

RP SEQUENCE OF 69-133 AND 312-329.

RX MEDLINE=71058486; PubMed=5538616;

RA Turner K.J., Cebr J.J.;

RT "Structure of heavy chain from strain 13 guinea pig

RT immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal

RT and hinge region cyanogen bromide fragments.";

RN Biochemistry 10:9-17(1971).

RN [4]

RP SEQUENCE OF 134-226.

RX MEDLINE=75036072; PubMed=4429665;

RA Tracey D.E., Cebr J.J.;

RT "Primary structure of the CH2 homology region from guinea pig IgG2

RT antibodies.";

RN Biochemistry 13:4796-4803(1974).

RN [5]

RP SEQUENCE OF 227-311.

RX MEDLINE=75036073; PubMed=4609467;

RA Trischmann T.M., Cebr J.J.;

RT "Primary structure of the CH3 homology region from guinea pig IgG2

RT antibodies.";

RN Biochemistry 13:4804-4811(1974).

RN [6]

RP DISULFIDE BONDS.

RX MEDLINE=71058474; PubMed=4922544;

RA Oliveira B., Lamm M.B.;

RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";

RN Biochemistry 10:26-31(1971).

CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN

CC 13 INBRED GUINEA PIGS.

DR PIR: A94553; G2GP.

DR HSSP: P01842; 7FAB.

DR InterPro: IPR007110; Ig-like.

DR InterPro: IPR003597; Ig\_c1.

DR InterPro: IPR003006; Ig\_MHC.

DR Pfam: PF00047; Ig; 2.

DR SMART: SM00407; IgG1; 2.

DR PROSITE: PS00835; IG\_LIKE; 3.

DR PROSITE: PS00290; IG\_MHC; 1.

KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.

FT NON\_TER 1

FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).

FT DISULFID 28 79

FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).

FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).

FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).

FT DISULFID 142 202

FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).

FT DISULFID 248 308

SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;

Query Match 30.5%; Score 887; DB 1; Length 329;

Best Local Similarity 65.9%; Pred. No. 5.8e-56;

Matches 166; Conservative 29; Mismatches 49; Indels 8; Gaps 3;

Qy 295 YPEEDTEGDK-VRAIRTRVWKFPPTKARPHCTPCPAPEALGAPSVFLPPKPKDTLMISR 353

Db 83 HPASSTKVDKTVETPIRT-----PZPBCTCPKPPENLGSPSVFIFPPKPKDTLMISLT 137

Qy 354 PEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVWSVLTVLHQDWLN 413

Db 138 PRVTCVVVDVSDPEVQFTWYINNEQVARTAPPRLREQQFNSTIRVWSTLPITHQDWLRG 197

Qy 414 KEYCKVSKNALKPVIETKISKAKGQPREPOVYITLPPSEEMTKNQVSLTCLAVKGFPVPSD 473

Db 198 KEFKCKVHNAKPAPIETKISKARGQPLEPKPVYTMGPPEELSSRSVSLTCMINGFPAD 257

Qy 474 IAVESNGOP--ENNYKTTTPVLDSGCSFFLYSKLTVDKSRWQGNVFCVSVNHEALHN 531

Db 258 IAVESNGOP--ENNYKTTTPVLDSGCSFFLYSKLTVDKSRWQGNVFCVSVNHEALHN 317

Qy 532 HYTKSKLSLSPG 543

Db 318 HVTQKISRSPG 329

RESULT 8

GC3\_MOUSE STANDARD; PRT; 329 AA.

ID - GC3\_MOUSE

AC P22436;

DT 01-AUG-1991 (Rel. 19, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Ig gamma-3 chain C region, secreted form.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=85027161; PubMed=6092053;

RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,

RA Tucker P.M., Blattner F.R.;

RT "Structural analysis of the murine IgG3 constant region gene.";

RN EMBO J. 3:2041-2046(1984).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



```

CC EMBL; J00451; -; NOT_ANNOTATED_CDS.
CC PIR; B02156; G3MSC.
CC HSSP; P01857; 1FC1.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003597; Ig_cl.
CC InterPro; IPR003006; Ig_MHC.
CC Pfam; PF00047; Ig; 3.
CC SMART; SM00407; IGC1; 2.
CC PROSITE; PS00835; IG_LIKE; 3.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CHI.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT TRANSMEM 346 362
FT DOMAIN 363 398
FT CONFLICT 333 333 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;

Query Match 29.0%; Score 843; DB 1; Length 329;
Best Local Similarity 65.5%; Pred. No. 7.8e-53;
Matches 152; Conservative 32; Mismatches 46; Indels 2; Gaps 1;

QY 315 PTKARHTCP--PCPAPEALGAPSVLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKF 372
DB 98 PRIPKSTPPGSSCPGNTILGSPSVFIFFPKPKDALMISLTPEKVTCTVVDVSEDDPDVHV 157

QY 373 NMVVDGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGKCKVSNKALPAPIERT 432
DB 158 SWFVDNKEVHTAWTPREAQYNSTFRVVSALPIQHODWNRGKFKCKVNNKALPAPIERT 217

QY 433 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 492
DB 218 ISKPKGRAQTPQVYITPPPREQMSKKKSVSLTCLVNTFFSEALSVENRGELEQDYKNT 277

QY 493 PVLDSGSPFLYSKLTVDKSRWQGNVFCVSMVHEALHNHYTQKSLSLSPGK 544
DB 278 PILDSDGTFLYSKLTVDTSMLQGEIFTCSVVHEALHNHHTQKNLSRSPGK 329

RESULT 9
GC3M_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DC 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-3 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Weis J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=8401483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene
segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J00451; AAB59655.1; -.
CC EMBL; V01526; CAA24767.1; ALT_SEQ.
CC PIR; A02156; G3MSM.
CC HSSP; P01857; 1FC1.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003597; Ig_cl.
CC InterPro; IPR003006; Ig_MHC.
CC Pfam; PF00047; Ig; 3.
CC SMART; SM00407; IGC1; 2.
CC PROSITE; PS00835; IG_LIKE; 3.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CHI.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT TRANSMEM 346 362
FT DOMAIN 363 398
FT CONFLICT 333 333 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA; 43929 MW; CF7264B50A41B95 CRC64;

Query Match 28.6%; Score 832; DB 1; Length 398;
Best Local Similarity 65.2%; Pred. No. 6e-52;
Matches 150; Conservative 32; Mismatches 46; Indels 2; Gaps 1;

QY 315 PTKARHTCP--PCPAPEALGAPSVLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKF 372
DB 98 PRIPKSTPPGSSCPGNTILGSPSVFIFFPKPKDALMISLTPEKVTCTVVDVSEDDPDVHV 157

QY 373 NMVVDGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGKCKVSNKALPAPIERT 432
DB 158 SWFVDNKEVHTAWTPREAQYNSTFRVVSALPIQHODWNRGKFKCKVNNKALPAPIERT 217

QY 433 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 492
DB 218 ISKPKGRAQTPQVYITPPPREQMSKKKSVSLTCLVNTFFSEALSVENRGELEQDYKNT 277

QY 493 PVLDSGSPFLYSKLTVDKSRWQGNVFCVSMVHEALHNHYTQKSLSLSP 542
DB 278 PILDSDGTFLYSKLTVDTSMLQGEIFTCSVVHEALHNHHTQKNLSRSP 327

RESULT 10
GC1_RAT STANDARD; PRT; 326 AA.
AC P20759;
DC 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-1 chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR; PS0017; PS0017.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003006; Ig_MHC.

```

DR Pfam; PF00047; ig; 3.  
DR SMART; SM00407; Igel; 2.  
DR PROSITE; PS50835; IG LIKE; 3.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.  
FT NON\_TER 1 1  
FT DOMAIN 1 97  
FT DOMAIN 98 112  
FT DOMAIN 113 219  
FT DOMAIN 220 326  
FT DISULFID 27 82  
FT DISULFID 102 102  
FT DISULFID 106 106  
FT DISULFID 109 109  
FT DISULFID 111 111  
FT DISULFID 140 200  
FT DISULFID 246 304  
FT CARBOHYD 176 176  
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 28.6%; Score 831; DB 1; Length 326;  
Best Local Similarity 44.8%; Pred. No. 5.5e-52;  
Matches 165; Conservative 50; Mismatches 77; Indels 76; Gaps 6;  
QY 195 LOENSLYT---IPKGF-----GSHLLPFAFLGNPWLNCCEILYFRWL 236  
Db 17 LKSNMVLGLCVKGYPEPVTWNSGALSSGVHTFAVLQSG-----60  
QY 237 QDNAENVYKQGVVDKAMTSNVAQCDNSDKFPVKYKPKGPTLGDDEGTDLYDYP 296  
Db 61 -----LYLTSSVTSSWPSQVTCNVA-----HP 87  
QY 297 EEDTEGDKVTRATRVVKEPTKARPHTCPPAPALGAPSVLFPKPKDTLMISRTPEV 356  
Db 88 ASSTKVDK-----KIVPRNCQGDCKPCICTGS-EVSSVFIPFPKPKDVLITTPKV 138  
QY 357 TCVVVDVSHEDPEVKFNMYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEY 416  
Db 139 TCVVVDISQDDPEVHFSFVDDVEVHTAQTTPPEQFNSTFRSVELPILHQDWLNGRTF 198  
QY 417 KCKVSNKALPVIETISKAKGQPREPOVYTLPPSRREMTKNQVSLCFLVKGYFSPDIAV 476  
Db 199 RCKVTSAAFPSPVETISKPEGRQVPHVYVMTSPKEEMTQNEVSIITCMVKGFYPPDIY 258  
QY 477 EWESNGQENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQK 536  
Db 259 EQWMQSQENYKNTPTMTDIDGSIYFLYSKLVNKKRQCGNTFTCSVLHEGLNHRHTEK 318  
QY 537 SLSLSPGK 544  
Db 319 SLSHSPGK 326

RESULT 11  
GCC RAT ID GCC RAT STANDARD; PRT; 333 AA.  
AC P20761;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Ig gamma-2B chain C region.  
OS Rattus norvegicus (Rat).  
OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89232738; PubMed=3149946;  
RA Brueggemann M.;  
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family."  
RL Gene 74:473-482(1988).  
CC !- SIMILARITY: Contains 3 immunoglobulin-like domains.

DR HSP; P01842; 7FAB.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003597; Ig\_C1.  
DR InterPro; IPR003006; Ig\_MHC.  
DR Pfam; PF00047; ig; 3.  
DR SMART; SM00407; Igel; 2.  
DR PROSITE; PS50835; IG LIKE; 3.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region; Repeat.  
FT NON\_TER 1 1  
FT DOMAIN 6 96  
FT DOMAIN 124 223  
FT DOMAIN 232 328  
FT DISULFID 15 15  
FT DISULFID 27 80  
FT DISULFID 106 106  
FT DISULFID 109 109  
FT DISULFID 112 112  
FT DISULFID 115 115  
FT DISULFID 147 207  
FT DISULFID 253 311  
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;  
Query Match 28.5%; Score 828; DB 1; Length 333;  
Best Local Similarity 50.1%; Pred. No. 9.2e-52;  
Matches 169; Conservative 42; Mismatches 92; Indels 34; Gaps 6;  
QY 237 QDNAENVYKQGVVDKAMTSNVAQCDNSDKFP-----VYKYPG---K 278  
Db 2 QTTAPSVYPLAPGCG---DTTSSTVTLGLVKGYPEPVTWNSGALSSDVFHTFPVQLS 59  
QY 279 GCPTLGDGDTLDYD-----YYPEEDTEGDKVTRATRVVKEPTKARPHTCPPCP 327  
Db 60 GLYTLTSSVTSWPSQVTCNVAHPASSTKVKKVERNGGIGHKCTCP---TCHKCP 116  
QY 328 APEALGAPSVLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTX 387  
Db 117 VPELLGGSPVIFPPKPKDILLISQNAKVTCCVVVDVSEEPDVQFSWVNNVEVHTAQ 176  
QY 388 PREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVIETISKAKGQPREPOVYT 447  
Db 177 PREEQYNSTFRVSALPTIQQDWMNSGKPKCKVNNKALPVIETISKPKGLVRKPKQVY 236  
QY 448 LPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQENNYKTPPVLDSDGSFFLYSKL 507  
Db 237 MGPTPEQITETQVSLTCLTSGLFNDIGVETNSGHIKVKYKNTPEPWDSGDSGFFMYSKL 296  
QY 508 TVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 544  
Db 297 NVESRSMDSRAPFVCSVVHGLNHRHVEKSISSRPFGK 333

RESULT 12  
GCC RAT ID GCC RAT STANDARD; PRT; 329 AA.  
AC P20762;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig gamma-2C chain C region.  
OS Rattus norvegicus (Rat).  
OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88166903; PubMed=3127222;  
RA Brueggemann M.; Delmastro-Galfire P.; Waldmann H.; Calabi F.;  
RT "Sequence of a rat immunoglobulin gamma 2C heavy chain constant  
region cDNA: extensive homology to mouse gamma 3."  
RL Eur. J. Immunol. 18:317-319(1988).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration



```
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN) .
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN) .
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN) .
FT DISULFID 138 198
FT CARBOHYD 174 174 /FTIG=CAR_000055.
FT DISULFID 244 302 REMOVED POST-TRANSLATIONALLY.
FT MOD RES 324 324 N -> D (IN REF. 3).
FT CONFLICT 276 276 N -> D (IN REF. 3).
FT CONFLICT 278 31.
SQ SEQUENCE 324 AA; 35704 MW; A338812FD3D1F2C93 CRC64;

Query Match 28.2%; Score 820.5; DB 1; Length 324;
Best Local Similarity 57.9%; Pred. No. 3e-51;
Matches 146; Conservative 43; Mismatches 48; Indels 15; Gaps 3;

Qy 295 YPEEDTEGDKVRAIRTVVKEPTKARPHTCPP--CPAPEALGAPSVFLPPPKDPTLMISR 352
Db 86 HPASSTKVDKKIVPRDC-----GCKPCICTVPE---VSSVFIFPPKPKDVLITL 132
Qy 353 TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPREOYNSTIRYVSQVLTVLHODWLN 412
Db 133 TPKVTCVVVDISKDPEVFQSFVDFVDDVEVHTAQTQPREEQFNSTFRSVSELPIMHODWLN 192
Qy 413 GREYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSREEMTKNOVSLTCLVKGFYPS 472
Db 193 GKEFKCVNNAFAFAPAEIKTISKGRKAPQVYIIPPKQMAKDKVSLCKMTDFPPE 252
Qy 473 DIAVEVNGSGQENNYKTPPPVLDSDGSFLLSKLTVDKSRWQQQNVFSCVHMPEALNHH 532
Db 253 DITVEWQMGQPAENYKNTQPIMTNGSYFVYSLKLVQKSNWEAGNTFTCSVLHGLNHH 312
Qy 533 YTKSLSLSPGK 544
Db 313 HTEKSLSHSPGK 324

RESULT 14
GCLM_MOUSE
ID GCLM_MOUSE STANDARD; PRT; 393 AA.
AC P01869;
DC 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig gamma-1 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";
RL Cell 18:559-568(1979).
RN [2]
RP SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
conserved transmembrane sequence and a 28-residue intracellular
domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
RN [3]
RP SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
```

```

[4]
SEQUENCE OF 1-44 FROM N.A.
MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
immunoglobulin gamma chains.";
Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
-1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=Membrane-bound;
IsoId=P01869-1; Sequence=Displayed;
Name=Secreted;
IsoId=P01868-1; Sequence=External;
Note=May be the major isoform;
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; V00793; CAA24172.1; -
EMBL; V00793; CAA24173.1; -
EMBL; V00793; CAA24174.1; -
PIR; B02159; GINSM
PDB; 1SC8; 23-MAR-99.
PDB; 1AB6; 18-MAR-98.
PDB; 1CL7; 12-JAN-00.
PDB; 1F11; 06-FEB-01.
PDB; 1F58; 29-DEC-99.
PDB; 1KCS; 24-JUL-02.
PDB; 1KCR; 11-MAY-02.
PDB; 25C8; 09-JUL-99.
MGD; MGI:96446; Igh-4.
InterPro; IPR007110; Ig-like.
InterPro; IPR003597; Ig cl.
InterPro; IPR003006; Ig_MHC.
Pfam; PF00047; Ig; 3.
SMART; SM00407; IGcl; 2.
PROSITE; PS00835; IG_LIKE; 3.
PROSITE; PS00290; IG_MHC; 1.
Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
Alternative splicing; Transmembrane; 3D-structure.
NON_TER 1
DOMAIN 1 97 CHI.
DOMAIN 98 110 HINGE.
DOMAIN 111 217 CH2.
DOMAIN 218 324 CH3.
DISULFID 27 82 INTERCHAIN (WITH A LIGHT CHAIN).
DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
DISULFID 138 198 N-LINKED (GLCNAC. . .).
CARBOHYD 174 174
DISULFID 244 302 POTENTIAL.
DISULFID 340 357 CYTOPLASMIC (POTENTIAL).
DOMAIN 358 393
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;

Query Match 28.1%; Score 815.5; DB 1; Length 393;
Best Local Similarity 57.8%; Pred. No. 8.9e-51;
Matches 145; Conservative 43; Mismatches 48; Indels 15; Gaps 3;

Qy 295 YPEEDTEGDKVRAIRTVVKEPTKARPHTCPP--CPAPEALGAPSVFLPPPKDPTLMISR 352
Db 86 HPASSTKVDKKIVPRDC-----GCKPCICTVPE---VSSVFIFPPKPKDVLITL 132
Qy 353 TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPREOYNSTIRYVSQVLTVLHODWLN 412
Db 133 TPKVTCVVVDISKDPEVFQSFVDFVDDVEVHTAQTQPREEQFNSTFRSVSELPIMHODWLN 192
```

```
QY 413 GKEYCKVSNKALPVPIEKTISKAKGQPREPOVYTLPPSRREEMTKNQVSLTCLVKGFYPS 472
DB 193 GKEFKRVNSAAPPAPRIEKTISKGRKPAQVYIIPPKQMAKDKVSLCTMTDFEPE 252
QY 473 DIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSPSCSMHEALHNH 532
DB 253 DITVEQMNGQPAENYKNTQPIMTNGSYFYVYSLKLVQKSNWEAGNTFTCSVLHREGLNHN 312
QY 533 YTKQSLSLSPG 543
DB 313 HTEKSLSHSPG 323

RESULT 15
GCA_RAT STANDARD; PRT; 322 AA.
AC P20760;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig gamma-2A chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_TaxID=10116;
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
CC -!- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M13804; AAA41376.1; ALT_INIT.
DR PIR; PS0019; PS0019.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; ig; 2.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Repeat.
FT NON_TER 1
FT DOMAIN 6 98 IG-LIKE 1.
FT DOMAIN 115 212 IG-LIKE 2.
FT DOMAIN 221 317 IG-LIKE 3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 136 196
FT DISULFID 242 300
FT CARBOHYD 172 172 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 322 AA; 35186 MW; E8EA136A9DE01EDB CRC64;

Query Match 27.9%; Score 812; DB 1; Length 322;
Best Local Similarity 44.3%; Pred. No. 1.2e-50;
Matches 163; Conservative 46; Mismatches 79; Indels 80; Gaps 6;

QY 195 LOENSLYLT----IPKGF-----GSHLLPFAFLHGNPMLCNCEILYFRWL 236
DB 17 LKSNMWTLCGLVKGYFPEPVTVTWSGALSSGVHTFPAVLQSG----- 60
```

```
QY 237 QDNAENYVWKQGVYKAMTSNVAQVQDNQSKFPVYKYFGKGPCTILGDEGDTLDLYDYP 296
DB 61 -----LYTLTSSVTPSSWSSQAVTCNVA-----HP 87
QY 297 BEDTEGDKVRAITRVVVFPTKARPHTCPPCPAPEALGAPSVFLPPPKPKDTLMISRTPEV 356
DB 88 ASSTKVDK-----KIVPRECNPCGCTGS-EVSSVFIPFPKTKDVLTLITLPKV 134
QY 357 TCVVVDVSHEDPEVKFNWYVDGVEVRNAAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEY 416
DB 135 TCVVVDISQNDPEVRFVSWFIDDDVEVHTAQTAAPEKQSNSTLRSSVSELPIVHRDLWNGKTF 194
QY 417 KCKVSNKALPVIIEKTIISKAKGQPREPOVYTLPPSRREEMTKNQVSLTCLVKGFYPSDIAV 476
DB 195 KCKVNSCAFPAPIEKSIKPEGTPRGQVYTMAPPKKEMTQSQVSIICWKGYFPDPIYT 254
QY 477 EWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSPSCSMHEALHNHYTQK 536
DB 255 EMKMGQGPQENYKNTPTPTMDTDGSGYFLYSLKLVKVKETMQQGNITFTCSVLHREGLNHNHTEK 314
QY 537 SLSLSPGK 544
DB 315 SLSHSPGK 322

Search completed: October 14, 2003, 06:31:08
Job time : 19.1814 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:16:09 ; Search time 57.1833 Seconds  
(without alignments)  
2454.925 Million cell updates/sec

Title: US-10-068-426-1

Perfect score: 2906

Sequence: 1 MPLLLLLLLLPLPHPHPIC.....MHEALNHYTQKLSLSPGK 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp.archaea.\*
- 2: sp.bacteria.\*
- 3: sp.fungi.\*
- 4: sp.human.\*
- 5: sp.invertebrate.\*
- 6: sp.mammal.\*
- 7: sp.mhc.\*
- 8: sp.organelle.\*
- 9: sp.phage.\*
- 10: sp.plant.\*
- 11: sp.podent.\*
- 12: sp.virus.\*
- 13: sp.vertebrate.\*
- 14: sp.unclassified.\*
- 15: sp.xvirus.\*
- 16: sp.bacteriap.\*
- 17: sp.archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1693	58.3	626	4	Q8N1F3
2	1687	58.1	626	4	Q8NG39
3	1224.5	42.1	701	4	Q96Q08
4	1207.5	41.6	471	4	Q8TC77
5	1137.5	39.1	473	4	Q8TC63
6	1137.5	39.1	521	4	Q8N4Y9
7	1126.5	38.8	509	4	Q8N1F7
8	1076.5	37.0	734	11	O35930
9	970	33.4	677	6	Q28256
10	891	30.7	337	6	Q95W34
11	838	28.8	469	11	Q8R3V9
12	835	28.7	463	11	Q93LC4
13	831	28.6	437	11	Q9R1A4
14	808	27.8	473	11	Q908L4
15	801	27.6	468	11	Q99L31
16	801	27.6	473	11	Q99L25

17	775	26.7	473	11	Q91Z05	Q91Z05 mus musculu
18	773	26.6	474	11	Q8R3H6	Q8R3H6 mus musculu
19	358	12.3	375	4	Q9BS21	Q9BS21 homo sapien
20	358	12.3	597	4	Q9BU10	Q9BU10 homo sapien
21	358	12.3	597	4	Q9BQB8	Q9BQB8 homo sapien
22	358	12.3	597	4	Q96BB9	Q96BB9 homo sapien
23	348	12.0	588	4	Q8WUX4	Q8WUX4 homo sapien
24	348	12.0	613	4	Q96EY0	Q96EY0 homo sapien
25	348	12.0	613	4	Q8WUK1	Q8WUK1 homo sapien
26	348	12.0	614	4	Q96GAG	Q96GAG homo sapien
27	348	12.0	618	4	Q96AA6	Q96AA6 homo sapien
28	342	11.8	613	11	Q8VCX7	Q8VCX7 mus musculu
29	302.5	10.4	342	11	Q91XL1	Q91XL1 mus musculu
30	301	10.4	347	4	Q8N4F5	Q8N4F5 homo sapien
31	293.5	10.1	1328	5	Q21043	Q21043 caenorhabdi
32	287	9.9	486	11	Q91Z07	Q91Z07 mus musculu
33	286	9.8	384	4	Q9UP60	Q9UP60 homo sapien
34	286	9.8	493	4	Q8NCL6	Q8NCL6 homo sapien
35	286	9.8	494	4	Q96K68	Q96K68 homo sapien
36	286	9.8	496	4	Q96KX8	Q96KX8 homo sapien
37	284	9.8	499	4	Q8N5K4	Q8N5K4 homo sapien
38	283	9.7	496	4	Q96DK0	Q96DK0 homo sapien
39	283	9.7	500	4	Q9BRV0	Q9BRV0 homo sapien
40	281	9.7	426	11	Q9DCD9	Q9DCD9 mus musculu
41	281	9.7	497	4	Q8WY24	Q8WY24 homo sapien
42	280	9.6	487	11	Q99KA4	Q99KA4 mus musculu
43	279	9.6	416	4	Q9NPP6	Q9NPP6 homo sapien
44	279	9.6	1527	5	Q9VZ24	Q9VZ24 drosophila
45	277.5	9.5	479	11	Q91WP5	Q91WP5 mus musculu

ALIGNMENTS

RESULT 1

Q8N1F3 ID Q8N1F3 PRELIMINARY; PRT; 626 AA.  
AC Q8N1F3;  
DT 01-OCT-2002 (TReMBLrel. 22, Created)  
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
DE Glycoprotein IB (Platelet), alpha polypeptide.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_taxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RA Strausberg R.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC027955; AAH27955.1; -.  
DR InterPro; IPR001611; LRR\_Cterm.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003591; LRR typ.  
DR InterPro; IPR002965; P-rich\_extensn.  
DR Pfam; PF00560; LRR; 1.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01462; LRRNT; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00127; PRICHEXTENS.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR typ; 5.  
DR PROSITE; PS0506; LRR\_TYPICAL; 1.  
SQ SEQUENCE 626 AA; 68969 MW; 1945FE52545DF573 CRC64;

Query Match 58.3%; Score 1693; DB 4; Length 626;

Best Local Similarity 87.4%; Pred.No. 1.1e-131;

Matches 333; Conservative 4; Mismatches 18; Indels 26; Gaps 5;

QY 1 MPLLLLLLLLPLPHPHPICVSKVASHLEVNCNRLTALPDPDKDTTILHSENLLY 60



```
Db 1 MLLLLLLLLLPSPLPHPICEVSKVASHLEVNCNRLTALPPDLKDTTILHLENLLY 60
QY 61 TFSLATLMPYTRLTQNLDRCELTKQVODGTLPVLTGLDLSHNQLOSLPLGQTLPALTV 120
Db 61 TFSLATLMPYTRLTQNLDRCELTKQVODGTLPVLTGLDLSHNQLOSLPLGQTLPALTV 120
QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLELP 180
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLELP 180
QY 181 AGLNGLENLDTLLLOENSlyTIPKFFGSHLLPFAFLHGNPMLCNCeILYFRWLQDNA 240
Db 181 AGLNGLENLDTLLLOENSlyTIPKFFGSHLLPFAFLHGNPMLCNCeILYFRWLQDNA 240
QY 241 ENVVYVKGQVDVKAMTSNVASVQCNSDKFPVYKPGKCGPTLGDGDTLDLYDYYPEEDT 300
Db 241 ENVVYVKGQVDVKAMTSNVASVQCNSDKFPVYKPGKCGPTLGDGDTLDLYDYYPEEDT 300
QY 301 EGDKVRATRTVVKPTKARPHTCP-----PCPA---PEALGAPSVFLPPPK 343
Db 301 EGDKVRATRTVVKPTKARPHTCP-----PCPA---PEALGAPSVFLPPPK 343
QY 344 --PKDTL-----MISRTPEVT 357
Db 359 WTPNFTLHMESITFSKTPKST 379

RESULT 2
Q8NG39 PRELIMINARY; PRT; 626 AA.
AC Q8NG39;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Platelet glycoprotein Ib alpha.
GN GPIBA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Matsumura Y., Murata M., Sugita K., Ikeda Y.;
RT "Identification of a novel point mutation in platelet glycoprotein
RT Iba, Gly co Ser at residue 233, in a Japanese family with platelet-
RT type von Willebrand disease.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB086948; BAC10305.1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003591; LRR_typ.
DR InterPro; IPR002965; P_Rich_extensn.
DR Pfam; PF00560; LRR; 1.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR PRINTS; PR01217; PRICHEXTENS.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 5.
DR PROSITE; PS50506; LRR_TYPICAL; 1.
SQ SEQUENCE 626 AA; 68985 MW; C7931FD07458B17P CRC64;
```

```
Query Match 58.1%; Score 1687; DB 4; Length 626;
Best Local Similarity 87.1%; Pred. No. 3.4e-131;
Matches 332; Conservative 4; Mismatches 19; Indels 26; Gaps 5;

QY 1 MLLLLLLLLLPSPLPHPICEVSKVASHLEVNCNRLTALPPDLKDTTILHLENLLY 60
Db 1 MLLLLLLLLLPSPLPHPICEVSKVASHLEVNCNRLTALPPDLKDTTILHLENLLY 60
```

```
QY 61 TFSLATLMPYTRLTQNLDRCELTKQVODGTLPVLTGLDLSHNQLOSLPLGQTLPALTV 120
Db 61 TFSLATLMPYTRLTQNLDRCELTKQVODGTLPVLTGLDLSHNQLOSLPLGQTLPALTV 120
QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLELP 180
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLELP 180
QY 181 AGLNGLENLDTLLLOENSlyTIPKFFGSHLLPFAFLHGNPMLCNCeILYFRWLQDNA 240
Db 181 AGLNGLENLDTLLLOENSlyTIPKFFGSHLLPFAFLHGNPMLCNCeILYFRWLQDNA 240
QY 241 ENVVYVKGQVDVKAMTSNVASVQCNSDKFPVYKPGKCGPTLGDGDTLDLYDYYPEEDT 300
Db 241 ENVVYVKGQVDVKAMTSNVASVQCNSDKFPVYKPGKCGPTLGDGDTLDLYDYYPEEDT 300
QY 301 EGDKVRATRTVVKPTKARPHTCP-----PCPA---PEALGAPSVFLPPPK 343
Db 301 EGDKVRATRTVVKPTKARPHTCP-----PCPA---PEALGAPSVFLPPPK 343
QY 344 --PKDTL-----MISRTPEVT 357
Db 359 WTPNFTLHMESITFSKTPKST 379

RESULT 3
Q96P08 PRELIMINARY; PRT; 701 AA.
AC Q96P08;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Factor VII active site mutant immunocognate.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hu Z., Garen A.;
RT "Targeting tissue factor on tumor vascular endothelial cells and tumor
RT cells for immunotherapy in mouse models of prostatic cancer.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF272774; AKS8686.1; -.
DR HSP; P00761; IAN1.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002383; GLA_blood.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00594; gla; 1.
DR Pfam; PF00047; ig; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00001; GLABLOOD.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00069; GLA; 1.
DR SMART; SM00407; IGc1; 1.
DR SMART; SM00020; Tryp_Sfc; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
```

```

DR PRG SITE; PS01187; EGF CA; 1.
DR PROSITE; PS00011; GLU CARBOXYLATION; 1.
DR PROSITE; PS00835; IG_LIKE; 2.
DR PROSITE; PS00290; IG_MHC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Hydrolase; Protease; Serine protease.
SQ SEQUENCE 701 AA; 71826 MW; 94AC5CEB42CC992F CRC64;

Query Match 42.1%; Score 1224.5; DB 4; Length 701;
Best Local Similarity 83.3%; Pred. No. 9.1e-93;
Matches 235; Conservative 5; Mismatches 23; Indels 19; Gaps 3;

QY 277 GKGCPTLGDEGDTLDYDYE-----EDTEGDKVRAITVVKPTKARP-----HT 322
DB 425 GGCATVGHGFYTRVSYQIEWLQKMRSEPRFGVLLRA-----PPFGSAEPKSCDKTHT 479

QY 323 CPPCPAPEALGAPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 382
DB 480 CPPCPAPELGGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 539

QY 383 NAKTKPREEQNSTYRVVSVLTVLHQDLNKGKYCKVKSNKALPVPPIETKISKAKQPRE 442
DB 540 NAKTKPREEQNSTYRVVSVLTVLHQDLNKGKYCKVKSNKALPAPIETKISKAKQPRE 599

QY 443 PQVTLTPSREEMTKNQVSLTCLVKGYFSPDSIAVEWESNGQPENNYKATTPPVLDSDGSFF 502
DB 600 PQVTLTPSRDELTKNQVSLTCLVKGYFSPDSIAVEWESNGQPENNYKATTPPVLDSDGSFF 659

QY 503 LYSKLTVDKSRWQQGNVFSQVMEALHNHYTQKSLSLSPGK 544
DB 660 LYSKLTVDKSRWQQGNVFSQVMEALHNHYTQKSLSLSPGK 701

RESULT 4
Q8TC77
ID Q8TC77 PRELIMINARY; PRT; 471 AA.
AC Q8TC77;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024289; AAH24289.1; -.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; IG_4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 2.
KW Hypothetical protein.
SQ SEQUENCE 471 AA; 51791 MW; 388F7F4CF588660E CRC64;

Query Match 41.6%; Score 1207.5; DB 4; Length 471;
Best Local Similarity 64.7%; Pred. No. 1.3e-91;
Matches 244; Conservative 19; Mismatches 43; Indels 71; Gaps 6;

QY 235 WLQDN---AENVYWKQGVVYKMTS-----NVASVQCDNSDKFPV---YKY 275
DB 99 YLQMSLRAEDTAVYICARDLRQLTSYWFYDLWGRGLTVTVSSASTKGPSVFFLAPSSKS 158

QY 276 PGKCPITLGDGSDTLDYDYEEDT-----EGDKVRAITVVKPKT 316

```

```

DB 159 TSGCTAALG---CLVKDYFPEPVTVMNSGALTSVGHVTPPAVLQSSGLYSLSVWTVPS 214
QY 317 KA-----RPHTPCPAPALGAPSVFLPPPKPKOT 347
DB 215 SSLGTQTYICNVNHPKSNITKVDKKVBPCKSDKTHTCPPCPAPELLGGLSPSVFLPPPKPKOT 274
QY 348 LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQNSTYRVVSVLTVLH 407
DB 275 LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQNSTYRVVSVLTVLH 334
QY 408 QDWLNGKEYKCKVSNKALPAPIETKISKAKQPREPOVYTLPPSREEMTKNQVSLTCLVK 467
DB 335 QDWLNGKEYKCKVSNKALPAPIETKISKAKQPREPOVYTLPPSREDELTKNQVSLTCLVK 394
QY 468 GFYPSDIAVEWESNGQPENNYKATTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSQVME 527
DB 395 GFYPSDIAVEWESNGQPENNYKATTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSQVME 454
QY 528 ALHNHYTQKSLSLSPGK 544
DB 455 ALHNHYTQKSLSLSPGK 471

RESULT 5
Q8TC63
ID Q8TC63 PRELIMINARY; PRT; 473 AA.
AC Q8TC63;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025985; AAH25985.1; -.
DR InterPro; IPR000923; BlueCu_1.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; IG_4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00196; COPPER_BLUE; 1.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 3.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51986 MW; E29920B09BA369F5 CRC64;

Query Match 39.1%; Score 1137.5; DB 4; Length 473;
Best Local Similarity 75.1%; Pred. No. 8.6e-86;
Matches 226; Conservative 17; Mismatches 42; Indels 16; Gaps 5;

QY 246 WKQGVYKAMTSNVAQVC--DNSDKFPVYKYPGKGCPTLGDEGDTLDYDYEEDTEGD 303
DB 187 WNSG----ALISGVHTPPPAVLQSSGLYSLSVTVPPSSSLGTKTYTCNVDHKP-SNTKVD 241
QY 304 KVRATRTVVVKFTKARPHTPCPAPALGAPSVFLPPPKPKDTLMISRTPEVTCVVVDV 363
DB 242 K-----RVESKYGP-PCPSCPAPFLGGPSVFLPPPKPKDTLMISRTPEVTCVVVDV 292
QY 364 SHEDPEVKFNWYVDGVEVHNAKTKPREEQNSTYRVVSVLTVLHQDLNKGKYCKVKSNK 423
DB 293 SOEQVQFNWYVDGVEVHNAKTKPREEQNSTYRVVSVLTVLHQDLNKGKYCKVKSNK 352
QY 424 ALPVPPIETKISKAKQPREPOVYTLPPSREEMTKNQVSLTCLVKGYFSPDSIAVEWESNGQ 483
DB 353 GLPSSIETKISKAKQPREPOVYTLPPSREEMTKNQVSLTCLVKGYFSPDSIAVEWESNGQ 412

```

Qy	484	PENNYKTTTPVLDSDG	SFFLYSKLTV	DKSRQOGNVF	SCSYMHEALNHV	TQKSLSLSPG	543
Db	413	PENNYKTTTPVLDSDG	SFFLYSRLTV	DKSRQOGNVF	SCSYMHEALNHV	TQKSLSLSLG	472
Qy	544	K	544				
		—					
Db	473	K	473				

RESULT 6	PRELIMINARY;	PRT;	521 AA.
Q8N4Y9	Q8N4Y9		
ID	Q8N4Y9		
AC	Q8N4Y9;		
DT	01-OCT-2002 (TrEMBLrel. 22, Created)		
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Hypothetical protein.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		

RP SEQUENCE FROM N.A.  
RC TISSUE=Primary B-Cells from Tonsils;  
RA Strausberg R;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR	EMBL; BC033178; AAH33178.1; -.
DR	InterPro; IPR007110; IG-1 LIKE.
DR	InterPro; IPR003597; IG-CL.
DR	InterPro; IPR003006; IG-MHC.
DR	InterPro; IPR003596; IG-V.
DR	Pfam; PF000047; IG; 4.
DR	SMART; SM00407; IGc1; 3.
DR	SMART; SM00406; IGv; 1.
DR	PROSITE; PS00835; IG- LIKE; 4.
DR	PROSITE; PS00290; IG-MHC; 2.
KW	Hypothetical protein.
SQ	SEQUENCE 521 AA; 57156 MW; 2AC7022E72D6CAA2 CRC64;

Query Match	39.1%	Score 1137.5;	DB 4;	Length 521;
Best Local Similarity	63.3%	Pred. NO. 9.9e-86;		
Matches 231;	Conservative	20;	Mismatches 57;	Indels 57;
Gaps				8;

Qy	194	LLOENSLY-----TIPKGFSGSHLLPFAFLHGNPWLNCBILYFRRWLQONNAENVYWK	247
Db	200	VLOSGLYLSLSSVTVPPSSLGT-----QYTCN-----VNHKFSNTKVBK	240
Qy	248	QGYDVVKAMTSNVA\$V-----QCDNSDKFPVKYPGKPTLGDEGDTLDLYVYPEED	299
Db	241	R-VELKTPLGDTHTTCRCRPEKSCUTPEPCRCPEP-KSCUT-----PPPC	285

[illegible]

QY	360	VVDYSHEDPEVKFNWVYDGVVEVHNAKTPREEQYNSTFRVVSVLTVLHQDWLNGKEYKCK	419
		:           :           :           :	
Db	337	VVDYSHEDPEVQFKWYDGVVEVHNAKTPREEQFNSTFRVVSVLTVLHQDWLNGKEYKCK	396

[illegible]

Qy	480	SNGQPENNYTTTPVLDSGSGFFLYSKLTVDKSRWQGQNVFSCSVNHEALHNHYTOKLS	539
		: :	
Db	457	SSGQPENNTTPPMLDSDGSGFFLYSKLTVDKSRWQGNIFSCSVNHEALHNRFTOKLS	516

Qy	540	LSPGK	544
Db	517	LSPGK	521

## RESULT 7

```

QBNF17
ID QBNF17 PRELIMINARY; PRT; 509 AA.
AC QBNF17
DT 01-OCT-2002 (trEMBLrel. 22, Created)
DT 01-OCT-2002 (trEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (trEMBLrel. 23, Last annotation update)
DE FJ000385 protein (fragment).
DE GN FLJ00385.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RC SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Jikuya H., Takano J., Kikuno R., Nagase T., Ohara O.;
RA "the nucleotide sequence of a cDNA clone isolated from human
RT spleen."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK090464; BAC03445.1; -.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG_c1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; ig; 3.
DR SMART; SM00407; IGc1; 3.
DR PROSITE; PS00835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 2.
DR NON TER
DR 1
SQ SEQUENCE 509 AA; 56111 MW; 089498D8076E863C CRC64;

```

Query Match 38.8%; Score 1126.5; DB 4; Length 509;  
Best Local Similarity 63.1%; Pred.No. 7.8e-85;

QY 194 LQENSLY-----TIPKGFSGSHLLPFAFLHGNPWLCNCEILYFRRWLQDNAENVVVMK 247  
          ::||: |||         ||: |||         :         :  
Db 119 VLSSGLYSLSVVTVPPSSLGT-----QTYTCN-----VNHKPSNTKVDK 159

**Qy**    248 QGVDMKTSNVASV-----QCDSNKFYVKYPGKGCPCTLGDEGDTLDLYDYPEED 299  
         :|::|       :||     |     |     |     |     |     |     |  
**Db**    160 R-VELKTPLGDTHTTCPRCEPKSCDTPPCPRCEP-KSCDT-----PPC 204

**Qy** 300 TEGDKVRA<sup>TR</sup>T<sup>VK</sup>FPT<sup>KAR</sup>PHTCPCPAEALGAPSVFLPPPKOTLMISRTPEVTCV 359  
:  
:  
:  
**Dd** 205 PRCEP<sup>PK</sup>SCDT-----PPCPRCPAPELLGGPSVFLPPPKOTLMISRTPEVTCV 255

Qy	360	WDYSHEDPEVKFNWYDGVGEVHNAKTPREEQVNSTYRVVSVLTVLHQDWLNGKEYCK	419
Db	256	WDYSHEDPEVQFKMYDGVGEVHNAKTPREEQFNSTFRVVSVLTVLHQDWLNGKEYCK	315

[illegible]

Qy	480	SNGOPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVPSCSWMEALHNHYTQKSLS	539
Db	376	SSGOPENNYKTTTPMLDSDGSFFLYSKLTVDKSRWQQGNIPSCSWMEALHNRFTQKSLS	435

Qy	540	LSP	542
Db	436	LSP	438

RESULT 8	
O35930	
ID O35930	PRELIMINARY; PRT: 734 AA.
AC O35930:	
DT 01-JAN-1998	(TRMELREL. 05, Created)
DT 01-JAN-1998	(TRMELREL. 05, Last sequence update)
DT 01-MAR-2003	(TRMELREL. 23, Last annotation update)
DE Platelet glycoprotein Ib-alpha.	
DE GPIbA.	
OS Mus musculus (Mouse).	

DT	01-DEC-2001	(TREMBlrel. 19, Created)
DT	01-DEC-2001	(TREMBlrel. 19, Last sequence update)

RESULT 10	
Q95M34	
ID	Q95M
AC	Q95M
DT	01-D
DT	01-D

```

DR 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Immunoglobulin gamma 1 heavy chain constant region (Fragment).
GN IGHCl.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
[1]
RN SEQUENCE FROM N.A.
RP
RA Wagner B.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=98383416; PubMed=9717671;
RA Wagner B., Overesch G., Sheoran A., Holmes M., Richards C.,
RA Leibold W., Radbruch A.;
RT "Organization of the equine immunoglobulin heavy chain constant region
RL genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";
RL Immunobiology 199,105-119(1998).
DR EMBL; AJ300675; CAC44624.1; -.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_2.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS0835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 2.
FT NON TER 1
SQ SEQUENCE 337 AA; 37438 MW; A60BF2B01DEFD1F6 CRC64;

Query Match 30.7%; Score 891; DB 6; Length 337;
Best Local Similarity 70.1%; Pred. No. 1.5e-65;
Matches 157; Conservative 34; Mismatches 31; Indels 2; Gaps 1;

QY 323 CPCCPAPEALGAPSVFLPFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVH 382
DB 114 CPKCPAPEALGGPSVFIFPPNPKDTLMITRTPEVTCVVDVSHEDPEVKFNWYVDGVEVH 173

QY 393 NAKTPREQVNSTYRVSVLTLVHODMLNGKEYCKKVSNAKALPVPTKTIKAKGQPRE 442
DB 174 TATTPKEQVNSTYRVSVLTLVHODMLNGKEYCKKVSNAKALPVPTKTIKAKGQPRE 233

QY 443 PQVYTLPSREEMTKNVSCLTLVKGFYPSDIAVWESNGQP--ENNYKTPPVLDSDGS 500
DB 234 PQVYLAHPHDSLSKSVTLCKYDFYPPEINEMQNGQPELETXYSTQAOQSDGS 293

QY 501 FFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 544
DB 294 YFLYSKLSVDRNRWQGGTFTTCGMHEALHNHYTKQNVKNPGK 337

RESULT 11
Q8R3V9 PRELIMINARY; PRT; 469 AA.
AC Q8R3V9;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024405; AAH24405.1; -.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 1.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE 463 AA; 51007 MW; EAA674C6BBC30783 CRC64;

Query Match 28.7%; Score 835; DB 11; Length 463;
Best Local Similarity 46.6%; Pred. No. 1e-60;
Matches 165; Conservative 54; Mismatches 83; Indels 52; Gaps 7;

QY 136 QENSlyT---IPKGFPGSHLLPFAFLHGNPWLNCNCEILYFRRWLQDNAENVYVMKQGV 252
DB 157 QTNQSVTLGCLVKGYFPEPVT-----VTWNSGSLSSGVHTFPAVLQ---SDLYTLSSSVTV 209

QY 253 KAMTSNVASVQCDNSDKFPVYKFGKCGPTLGDGDTLDYDYPEEDTEGDKVRAITRV 312
DB 216 PSSTWPSQTCNVA-----HPASSTKVDKTIKVPDRC- 247

QY 313 KFTPKARHTICPP--CPAPEALGAPSVFLPFPKPKDTLMISRTPEVTCVVDVSHEDPEV 370
DB 248 -----GCKPCICTVPE---VSSVFIPFPKPKDVLITLTLPKVTCTVVDVSHEDPEV 295

QY 371 KFNWYVDGVEVHNAKTKPREQYNSTYRVSVLTLVHODMLNGKEYCKKVSNAKALPVPIE 430
DB 296 QFSNFWDDVEVHTAQTKEPREQPNSTFRSVSELPIMHODMLNGKEYCKRVNSAFAFPAIE 355

QY 431 KTIKAKGQPREQVYTLPPSRBEMTKNQVSLTCLVKGFYPSDIAVWESNGQPKENYKT 490
DB 356 KTIKTKGRPKAPQVYTIIPPKQMAKDKVSLTCTMITDFFPEDITVEMWNGQPAENYKN 415

QY 491 TTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 544
DB 416 TQPIINDTGSFVYSKLVNQSNAEAGNTFTCSVLHGLNHHHTKLSLSPGK 469

RESULT 12
Q99LC4 PRELIMINARY; PRT; 463 AA.
ID Q99LC4;
AC Q99LC4;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Similar to RIKEN CDNA 1810060009 gene.
GN IGH-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC03435; AAH03435.1; -.
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 1.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE 463 AA; 51007 MW; EAA674C6BBC30783 CRC64;

Query Match 28.7%; Score 835; DB 11; Length 463;
Best Local Similarity 46.6%; Pred. No. 1e-60;
Matches 165; Conservative 54; Mismatches 83; Indels 52; Gaps 7;

QY 136 QENSlyT---IPKGFPGSHLLPFAFLHGNPWLNCNCEILYFRRWLQDNAENVYVMKQGV 252
DB 157 QTNQSVTLGCLVKGYFPEPVT-----VTWNSGSLSSGVHTFPAVLQ---SDLYTLSSSVTV 209

QY 253 KAMTSNVASVQCDNSDKFPVYKFGKCGPTLGDGDTLDYDYPEEDTEGDKVRAITRV 312
DB 216 PSSTWPSQTCNVA-----HPASSTKVDKTIKVPDRC- 247

QY 313 KFTPKARHTICPP--CPAPEALGAPSVFLPFPKPKDTLMISRTPEVTCVVDVSHEDPEV 370
DB 248 -----GCKPCICTVPE---VSSVFIPFPKPKDVLITLTLPKVTCTVVDVSHEDPEV 295

QY 371 KFNWYVDGVEVHNAKTKPREQYNSTYRVSVLTLVHODMLNGKEYCKKVSNAKALPVPIE 430
DB 296 QFSNFWDDVEVHTAQTKEPREQPNSTFRSVSELPIMHODMLNGKEYCKRVNSAFAFPAIE 355

QY 431 KTIKAKGQPREQVYTLPPSRBEMTKNQVSLTCLVKGFYPSDIAVWESNGQPKENYKT 490
DB 356 KTIKTKGRPKAPQVYTIIPPKQMAKDKVSLTCTMITDFFPEDITVEMWNGQPAENYKN 415

QY 491 TTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 544
DB 416 TQPIINDTGSFVYSKLVNQSNAEAGNTFTCSVLHGLNHHHTKLSLSPGK 469

```

```

Db 210 PSSWPSSETVTCNVA-----HPASSTKVKDKKIVPRDC- 241
Qy 313 KPFTKARPHTCPP--CPAPEALGAPSVFLPPPKPKDTLMISRTPTBVTCTVVDVSHEDPEV 370
Db 242 -----GCKPCICTVPE---VSSVFIPFPKPKDVLITLTPKVTCTVVDVSHEDPEV 289
Qy 371 KFNWYVDGVEVHNAKTFRPEEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALVPVPIE 430
Db 290 QFSWFVDVVEVHTAQTPREEQNSTFRVSELPIHQDNLNGKEFKCRVNSAAPAPIE 349
Qy 431 KTISKAGOPREPQVYTLPPSREMTKNQVSLTCLVKGYFSPSDIAVWESNGOPENNYKT 490
Db 350 KTISKTKGRPKAPQVYTIPPPKEQMAKDKVSLTCTMITDFEFDITVWQWNGOPAEYKN 409
Qy 491 TPVLDSDGSFFLYSKLTVDKSRWQOQNVFSCSVWHEALHNYHTQKSLSPGK 544
Db 410 TQPMIDTGSFYFVSKLVNOKSNWEAGNTFTCSVLHGLHNNHTKSLSHSPGK 463

RESULT 13
Q9R1A4 PRELIMINARY; PRT: 437 AA.
AC Q9R1A4
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
GN IGH-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Walde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFv).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152372; AAD04243.1; -.
DR HSP; P01842; 7FAB.
DR MGD; MGI:96446; IGH-4.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 1.
FT NON_TER 1
FT TER 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match 28.6%; Score 831; DB 11; Length 437;
Best Local Similarity 46.3%; Pred. No. 2e-60;
Matches 164; Conservative 55; Mismatches 83; Indels 52; Gaps 7;

Qy 196 QENSLYT---IPKGFSGSHLLPEAFLGNPWLNCNCILYFRMLQNAENVTYVWQGVVDV 252
Db 131 QTNSTMTGLCLVKGYFPEPT---VTWNSGSLSSGVHTFPVQLQ---SDLYTLSSVTY 183
Qy 253 KAMTSNVASVOCNSDKFPYKYPGKCPFLGDEGDTLDYVYYPEEDTSGDKVRAITRV 312
Db 184 PSSWPSSETVTCNVA-----HPASSTKVKDKKIVPRDC- 215
Qy 313 KPFTKARPHTCPP--CPAPEALGAPSVFLPPPKPKDTLMISRTPTBVTCTVVDVSHEDPEV 370
Db 216 -----GCKPCICTVPE---VSSVFIPFPKPKDVLITLTPKVTCTVVDVSHEDPEV 263
Qy 371 KFNWYVDGVEVHNAKTFRPEEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALVPVPIE 430
Db 264 QFSWFVDVVEVHTAQTPREEQNSTFRVSELPIHQDNLNGKEFKCRVNSAAPAPIE 323

```

```

Qy 431 KTISKAGOPREPQVYTLPPSREMTKNQVSLTCLVKGYFSPSDIAVWESNGOPENNYKT 490
Db 324 KTISKTKGRPKAPQVYTIPPPKEQMAKDKVSLTCTMITDFEFDITVWQWNGOPAEYKN 383
Qy 491 TPVLDSDGSFFLYSKLTVDKSRWQOQNVFSCSVWHEALHNYHTQKSLSPGK 544
Db 384 TQPMIDTGSFYFVSKLVNOKSNWEAGNTFTCSVLHGLHNNHTKSLSHSPGK 437

RESULT 14
Q9D8L4 PRELIMINARY; PRT: 473 AA.
AC Q9D8L4
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 181006009Rik protein.
GN IGH-1 OR 181006009RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Pancreas;
RA MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli K., Barch G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007918; BAB25349.1; -.
DR HSP; P01842; 7FAB.
DR MGD; MGI:96443; IGH-1.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 473 AA; 51699 MW; 9DED57A514475FBB CRC64;

Query Match 27.8%; Score 808; DB 11; Length 473;
Best Local Similarity 58.3%; Pred. No. 1.8e-58;
Matches 146; Conservative 39; Mismatches 59; Indels 8; Gaps 2;

Qy 295 YPEEDTSGDKVRAITRVVVKFPTKARP---HTCPCPAPEALGAPSVFLPPPKPKDTLMI 350
Db 224 HPASSTKVDK---KTEPRVPIQNFCPLKCEPCCAAPDLLGGPSVFIFPPKIDVLM 279
Qy 351 SRPEVTCTVVDVSHEDPEVKFNWYVDGVEVHNAKTFRPEEQNSTYRVVSVLTVLHODW 410
Db 280 SLSPMTCTVVDVSDSDPDVQISWFNVEVHTAQTPREEQNSTYRVVSVLTVLHODW 339
Qy 411 LNKGEYKCKVSNKALVPVPIEKTISKAGOPREPQVYTLPPSREMTKNQVSLTCLVKGY 470
Db 340 MSGKEFKCKVNNALPSPIEKTISKPRGPVRAPOVYVLPVPPPAEEMTKGFSLTGITGL 399

```





PA (GEMY) G

PA YY

PI Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
XX WPI; 2002-657537/70.  
DR N-PSDB; ABQ78661.  
XX  
XX New glycoprotein Ib alpha fusion polypeptides, useful for treating a  
PT disorder associated with platelet activation e.g. ischaemic heart  
PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
XX  
XX Claim 5; Page 2; 45pp; English.  
PS  
XX The present sequence represents a fusion protein of glycoprotein  
CC 1B-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion  
CC protein inhibits the adherence of platelets to leukocytes. The fusion  
CC polypeptide is useful for treating a disorder associated with platelet  
CC activation e.g. ischaemic heart disease, acute myocardial infarction,  
CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
CC unstable angina. It can also be used to treat vascular conditions  
CC associated with vascular inflammation, thrombosis, and  
CC angioplasty-related restenosis.  
XX  
XX Sequence 544 AA;  
SQ  
Query Match 100.0%; Score 2906; DB 23; Length 544;  
Best Local Similarity 100.0%; Pred. No. 4.9e-216;  
Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MPLLALLLPSPLPHPHICEVSKVASHLEVNCNRLTALPPDLPKDTTILHLSNLLY 60  
Db 1 MPLLALLLPSPLPHPHICEVSKVASHLEVNCNRLTALPPDLPKDTTILHLSNLLY 60  
QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQSLPLGQTLPALTV 120  
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQSLPLGQTLPALTV 120  
QY 121 LDVSPNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTLP 180  
Db 121 LDVSPNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTLP 180  
QY 181 AGLINGLENLDTLLQENSLYTIIPKFGFSGHLLPFAFLHGNPWLNCNCEIYFRRLQDNA 240  
Db 181 AGLINGLENLDTLLQENSLYTIIPKFGFSGHLLPFAFLHGNPWLNCNCEIYFRRLQDNA 240  
QY 241 ENVYVWKQGVVDVKAMTSNVASVQCDSNDFPVYKYPGKCGPTLGDGDTLDLYDYPEEDT 300  
Db 241 ENVYVWKQGVVDVKAMTSNVASVQCDSNDFPVYKYPGKCGPTLGDGDTLDLYDYPEEDT 300  
QY 301 EGDKVRATRTVVKFPTKARPHCTPCPCPAEALGAPSVFLFPKPKDTLMISRTPEVTCV 360  
Db 301 EGDKVRATRTVVKFPTKARPHCTPCPCPAEALGAPSVFLFPKPKDTLMISRTPEVTCV 360  
QY 361 VDVSHEDEPVKNWYDGVGVHNAKTTPREOYNSTYRVSVLTVLHQDLNGKEYKCKV 420  
Db 361 VDVSHEDEPVKNWYDGVGVHNAKTTPREOYNSTYRVSVLTVLHQDLNGKEYKCKV 420  
QY 421 SNKALVPVLEKTSKAGOPREPQVYTLPPSREMTKNQVSLTCLVKGYPSDIAVENES 480  
Db 421 SNKALVPVLEKTSKAGOPREPQVYTLPPSREMTKNQVSLTCLVKGYPSDIAVENES 480  
QY 481 NGOPENNYKTTTPVLSDSGSFFLYSKLTVDKSRWQQGNVFCGVMHEALHNNHYTKS-SL 540  
Db 481 NGOPENNYKTTTPVLSDSGSFFLYSKLTVDKSRWQQGNVFCGVMHEALHNNHYTKS-SL 540  
QY 541 SPQK 544  
Db 541 SPQK 544

RESULT 2  
ABB78235  
ID ABB78235 standard; Protein; 544 AA.  
XX  
AC ABB78235;

XX 25-NOV-2002 (first entry)  
DT Amino acid sequence of GPIb302/2A-Ig fusion protein.  
XX  
DE Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;  
XX leukocyte; platelet activation; ischaemic heart disease;  
KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
XX thrombosis; angioplasty; restenosis.  
XX Synthetic.  
OS  
XX Key Location/Qualifiers  
FH Misc-difference 500 /note="Ser encoded by CCC"  
FT  
XX WO200263003-A2.  
XX  
PD 15-AUG-2002.  
XX  
XX 06-FEB-2002; 2002WO-US03549.  
PF  
XX 06-FEB-2001; 2001US-266838P.  
PR  
XX (GEMY ) GENETICS INST LLC.  
PA  
XX Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
PI WPI; 2002-657537/70.  
XX N-PSDB; ABQ78662.  
DR  
XX New glycoprotein Ib alpha fusion polypeptides, useful for treating a  
PT disorder associated with platelet activation e.g. ischaemic heart  
PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
XX  
XX Claim 20; Page 3; 45pp; English.  
PS  
XX The present sequence represents a fusion protein of glycoprotein  
CC 1B-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion  
CC protein inhibits the adherence of platelets to leukocytes. The fusion  
CC polypeptide is useful for treating a disorder associated with platelet  
CC activation e.g. ischaemic heart disease, acute myocardial infarction,  
CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
CC unstable angina. It can also be used to treat vascular conditions  
CC associated with vascular inflammation, thrombosis, and  
CC angioplasty-related restenosis.  
XX  
XX Sequence 544 AA;  
SQ  
Query Match 99.6%; Score 2894; DB 23; Length 544;  
Best Local Similarity 99.6%; Pred. No. 4.1e-215;  
Matches 542; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MPLLALLLPSPLPHPHICEVSKVASHLEVNCNRLTALPPDLPKDTTILHLSNLLY 60  
Db 1 MPLLALLLPSPLPHPHICEVSKVASHLEVNCNRLTALPPDLPKDTTILHLSNLLY 60  
QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQSLPLGQTLPALTV 120  
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQSLPLGQTLPALTV 120  
QY 121 LDVSPNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTLP 180  
Db 121 LDVSPNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTLP 180  
QY 181 AGLINGLENLDTLLQENSLYTIIPKFGFSGHLLPFAFLHGNPWLNCNCEIYFRRLQDNA 240  
Db 181 AGLINGLENLDTLLQENSLYTIIPKFGFSGHLLPFAFLHGNPWLNCNCEIYFRRLQDNA 240  
QY 241 ENVYVWKQGVVDVKAMTSNVASVQCDSNDFPVYKYPGKCGPTLGDGDTLDLYDYPEEDT 300  
Db 241 ENVYVWKQGVVDVKAMTSNVASVQCDSNDFPVYKYPGKCGPTLGDGDTLDLYDYPEEDT 300



PD 15-AUG-2002.  
 XX  
 PF 06-FEB-2002; 2002WO-USC3549.  
 XX  
 PR 06-FEB-2001; 2001US-266838P.  
 XX  
 PA (GEM) ) GENETICS INST LLC.  
 XX  
 PI Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
 XX  
 DR WPI; 2002-657537/70.  
 DR N-PSDB; ABQ78664.  
 XX  
 XX New glycoprotein Ib alpha fusion polypeptides, useful for treating a  
 PT disorder associated with platelet activation e.g. ischaemic heart  
 PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
 XX  
 PS Claim 20; Page 3; 45pp; English.  
 XX  
 CC The present sequence represents a fusion protein of glycoprotein  
 CC IB-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion  
 CC protein inhibits the adherence of platelets to leukocytes. The fusion  
 CC polypeptide is useful for treating a disorder associated with platelet  
 CC activation e.g. ischaemic heart disease, acute myocardial infarction,  
 CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
 CC unstable angina. It can also be used to treat vascular conditions  
 CC associated with vascular inflammation, thrombosis, and  
 CC angioplasty-related restenosis.  
 XX  
 SQ Sequence 531 AA;  
 Query Match 97.2%; Score 2825.5; DB 23; Length 531;  
 Best Local Similarity 97.6%; Pred. No. 7.9e-210;  
 Matches 531; Conservative 0; Mismatches 0; Indels 13; Gaps 1;  
 QY 1 MPLLLLLLLLPSPLPHPHIPICEVSKVASHLEWNCCKRNLTALPPDLPKDTTILHLSNLLY 60  
 DB 1 MPLLLLLLLLPSPLPHPHIPICEVSKVASHLEWNCCKRNLTALPPDLPKDTTILHLSNLLY 60  
 QY 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDTGLPVLTGLDLSHNOQLSLPLGOTLPALT 120  
 DB 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDTGLPVLTGLDLSHNOQLSLPLGOTLPALT 120  
 QY 121 LDVSNFRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180  
 DB 121 LDVSNFRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180  
 QY 181 AGLNGLENLDTLLQENSLYTIIPKFFGSHLLPFAFLHGNPWLNCCELLYFRRWLQDNA 240  
 DB 181 AGLNGLENLDTLLQENSLYTIIPKFFGSHLLPFAFLHGNPWLNCCELLYFRRWLQDNA 240  
 QY 241 ENVYVWKQGVYDVKAMTSNVASVQDNDKFPVYKPGKCGPTLGDGDDTLDLYYPEEDT 300  
 DB 241 ENVYVWKQGVYDVKAMTSNVASVQDNDKFPVYKPGKCGPTLGDGDDTLDLYYPEEDT 300  
 QY 301 EGDKVRATRIWVFPTTKARPHTCPPCPAPALGAPSVFLFPKPKDTLMISRTPEVTCV 360  
 DB 301 EGDKV-----RPHTCPPCPAPALGAPSVFLFPKPKDTLMISRTPEVTCV 347  
 QY 361 VDVSHDEPVKFNWYDGVGVHNAKTPREQYNSTRVVSVLTVLHQDLNGKEYKCV 420  
 DB 348 VDVSHDEPVKFNWYDGVGVHNAKTPREQYNSTRVVSVLTVLHQDLNGKEYKCV 407  
 QY 421 SNKALPVPIDKTSKAGQPREPOVYTLPPSREMTKNQVSLTCLVKGFYFSPDIAVWES 480  
 DB 408 SNKALPVPIDKTSKAGQPREFOVYTLPPSREMTKNQVSLTCLVKGFYFSPDIAVWES 467  
 QY 481 NGQPENNYKTTTPPVLSDGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSL 540  
 DB 468 NGQPENNYKTTTPPVLSDGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSL 527  
 QY 541 SPKG 544  
 |||||

DB 528 SPKG 531  
 RESULT 5  
 ABB78239  
 ID ABB78239 standard; Protein; 531 AA.  
 XX  
 AC ABB78239;  
 XX  
 DT 25-NOV-2002 (first entry)  
 XX  
 DE Amino acid sequence of GPIb290/1A-Ig fusion protein.  
 XX  
 KW Glycoprotein IB-alpha; GPIb; immunoglobulin; Ig; platelet adherence;  
 KW leukocyte; platelet activation; ischaemic heart disease;  
 KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
 KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
 KW thrombosis; angioplasty; restenosis.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 253 /note= "Val encoded by GCG"  
 FT Misc-difference 487 /note= "Ser encoded by CCC"  
 XX  
 PN WO200263003-A2.  
 XX  
 PD 15-AUG-2002.  
 XX  
 PF 06-FEB-2002; 2002WO-USC3549.  
 XX  
 PR 06-FEB-2001; 2001US-266838P.  
 XX  
 PA (GEM) ) GENETICS INST LLC.  
 XX  
 PI Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
 XX  
 DR WPI; 2002-657537/70.  
 DR N-PSDB; ABQ78666.  
 XX  
 PT New glycoprotein Ib alpha fusion polypeptides, useful for treating a  
 PT disorder associated with platelet activation e.g. ischaemic heart  
 PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
 XX  
 PS Claim 20; Page 4; 45pp; English.  
 XX  
 CC The present sequence represents a fusion protein of glycoprotein  
 CC IB-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion  
 CC protein inhibits the adherence of platelets to leukocytes. The fusion  
 CC polypeptide is useful for treating a disorder associated with platelet  
 CC activation e.g. ischaemic heart disease, acute myocardial infarction,  
 CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
 CC unstable angina. It can also be used to treat vascular conditions  
 CC associated with vascular inflammation, thrombosis, and  
 CC angioplasty-related restenosis.  
 XX  
 SQ Sequence 531 AA;  
 Query Match 97.0%; Score 2818.5; DB 23; Length 531;  
 Best Local Similarity 97.4%; Pred. No. 2.7e-209;  
 Matches 530; Conservative 0; Mismatches 1; Indels 13; Gaps 1;  
 QY 1 MPLLLLLLLLPSPLPHPHIPICEVSKVASHLEWNCCKRNLTALPPDLPKDTTILHLSNLLY 60  
 DB 1 MPLLLLLLLLPSPLPHPHIPICEVSKVASHLEWNCCKRNLTALPPDLPKDTTILHLSNLLY 60  
 QY 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDTGLPVLTGLDLSHNOQLSLPLGOTLPALT 120  
 DB 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDTGLPVLTGLDLSHNOQLSLPLGOTLPALT 120  
 QY 121 LDVSNFRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180

Db 121 LDVSNRLTSLPLGALRGGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180  
 Qy 181 AGLNGLENLDTLLLOENSLYTIKGFSGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240  
 Db 181 AGLNGLENLDTLLLOENSLYTIKGFSGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240  
 Qy 241 ENVYVWKQGVVDKAMTSNVASVQCDSNDFPVYKYPGKGCPTLGDGDTLDLYYPEDT 300  
 Db 241 ENVYVWKQGVVDKAMTSNVASVQCDSNDFPVYKYPGKGCPTLGDGDTLDLYYPEDT 300  
 Qy 301 EGDKVRATRVVVKFPTKARPHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCW 360  
 Db 301 EGDKV-----RPHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCW 347  
 Qy 361 VDVSHEDEPKVFNWYVDGVEVHNATKPREEOYNSTYRVSVLTVLHODWLNGKEYKCKV 420  
 Db 348 VDVSHEDEPKVFNWYVDGVEVHNATKPREEOYNSTYRVSVLTVLHODWLNGKEYKCKV 407  
 Qy 421 SNKALPVPIEKTISKAGOPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVEMES 480  
 Db 408 SNKALPVPIEKTISKAGOPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVEMES 467  
 Qy 481 NGQPNKYKTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKSLSL 540  
 Db 468 NGQPNKYKTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKSLSL 527  
 Qy 541 SPKG 544  
 Db 528 SPKG 531

## RESULT 6

ABB78238  
 ID ABB78238 standard; Protein; 531 AA.

AC ABB78238;  
 XX

DT 25-NOV-2002 (first entry)  
 XX

DE Amino acid sequence of GPlb290/2V-Ig fusion protein.

XX Glycoprotein 1B-alpha; GPlb; immunoglobulin; Ig; platelet adherence;  
 KW leukocyte; platelet activation; ischaemic heart disease;  
 KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
 KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
 KW thrombosis; angioplasty; restenosis.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 487 /note= "Ser encoded by CCC"  
 FT

XX WO200263003-A2.

XX 15-AUG-2002.

XX 06-FEB-2002; 2002WO-US03549.

XX 06-FEB-2001; 2001US-266838P.

XX (GEMY ) GENETICS INST LLC.

XX Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;

XX WPI; 2002-657537/70.

XX N-PSDB; ABQ78665.

XX New glycoprotein 1b alpha fusion polypeptides, useful for treating a  
 PT disorder associated with platelet activation e.g. ischaemic heart  
 disease, stroke, venous or arterial thrombosis or atherosclerosis -

XX

Claim 20; Page 3-4; 45pp; English.

CC The present sequence represents a fusion protein of glycoprotein  
 CC 1B-alpha (GPlb) and an immunoglobulin (Ig) polypeptide. The fusion  
 CC protein inhibits the adherence of platelets to leukocytes. The fusion  
 CC polypeptide is useful for treating a disorder associated with platelet  
 CC activation e.g. ischaemic heart disease, acute myocardial infarction,  
 CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
 CC unstable angina. It can also be used to treat vascular conditions  
 CC associated with vascular inflammation, thrombosis, and  
 CC angioplasty-related restenosis.

SQ Sequence 531 AA;

Query Match 96.8%; Score 2812.5; DB 23; Length 531;  
 Best Local Similarity 97.2%; Pred. No. 8e-209; 1; Indels 13; Gaps 1;  
 Matches 529; Conservative 1; Mismatches 1;

Qy 1 MPELLLLLLLPSPLPHPICEVSKVASHLEVNCDEKNTALPPDPKDTTILHLSENLLY 60

Db 1 MPELLLLLLLPSPLPHPICEVSKVASHLEVNCDEKNTALPPDPKDTTILHLSENLLY 60

Qy 61 TFSLATIMPYTRLTQNLDRCELTKLQVDGTLPLVIGTLDLSHNQLOSLPLLGOTLPALT 120

Db 61 TFSLATIMPYTRLTQNLDRCELTKLQVDGTLPLVIGTLDLSHNQLOSLPLLGOTLPALT 120

Qy 121 LDVSNRLTSLPLGALRGGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180

Db 121 LDVSNRLTSLPLGALRGGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180

Qy 181 AGLNGLENLDTLLLOENSLYTIKGFSGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240

Db 181 AGLNGLENLDTLLLOENSLYTIKGFSGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240

Qy 241 ENVYVWKQGVVDKAMTSNVASVQCDSNDFPVYKYPGKGCPTLGDGDTLDLYYPEDT 300

Db 241 ENVYVWKQGVVDKAMTSNVASVQCDSNDFPVYKYPGKGCPTLGDGDTLDLYYPEDT 300

Qy 301 EGDKVRATRVVVKFPTKARPHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 360

Db 301 EGDKV-----RPHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 347

Qy 361 VDVSHEDEPKVFNWYVDGVEVHNATKPREEOYNSTYRVSVLTVLHODWLNGKEYKCKV 420

Db 348 VDVSHEDEPKVFNWYVDGVEVHNATKPREEOYNSTYRVSVLTVLHODWLNGKEYKCKV 407

Qy 421 SNKALPVPIEKTISKAGOPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVEMES 480

Db 408 SNKALPVPIEKTISKAGOPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVEMES 467

Qy 481 NGQPNKYKTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKSLSL 540

Db 468 NGQPNKYKTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNYTKSLSL 527

Qy 541 SPKG 544

Db 528 SPKG 531

## RESULT 7

AAY49933

ID AAY49933 standard; Protein; 562 AA.

XX AAY49933;

XX 01-FEB-2000 (first entry)

XX Human glycoprotein 1b/mouse IgG1Fc chimeric protein.

KW Glycoprotein 1b; glycoallidin; detection; antithrombotic; binding;

KW von Willebrand factor; botrocetin; chimeric protein; immunoglobulin;

KW thrombotic disease; diagnosis; cardiac infarction; cerebral embolism.

XX





CC for the detection of glykocallidin as a means of thrombotic disease  
CC diagnosis e.g. for cardiac infarction and cerebral embolism, and also  
CC for screening substances with anti-thrombotic activity for the  
CC prevention and treatment of thrombotic diseases. The method is direct,  
CC convenient and quantitative, with reproducibility, and there is no need  
CC to construct a monoclonal antibody for the assay. The present sequence  
CC represents a human glycoprotein Ib/mouse immunoglobulin gamma 2a Fc  
CC chimeric protein from the present invention.  
XX  
XX  
SQ Sequence 568 AA;  
  
Query Match 84.6%; Score 2459; DB 21; Length 568;  
Best Local Similarity 81.6%; Pred. No. 1.9e-181;  
Matches 465; Conservative 28; Mismatches 49; Indels 28; Gaps 3;  
  
Qy 1 MPLLLLLLLLPSPHPHPICEVSKVASHLEVNCCKRNLTPDLPKDTTILHLENLY 60  
Db 1 MPLLLLLLLLPSPHPHPICEVSKVASHLEVNCCKRNLTPDLPKDTTILHLENLY 60  
  
Qy 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQQLSLPLGQTLPALTV 120  
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQQLSLPLGQTLPALTV 120  
  
Qy 121 LDVSFNLRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKSLANNLTLP 180  
Db 121 LDVSFNLRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKSLANNLTLP 180  
  
Qy 181 AGLLNGLENDTLLOENSLYTIKPGFSGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240  
Db 181 AGLLNGLENDTLLOENSLYTIKPGFSGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240  
  
Qy 241 ENVVWVKQGVVDVKAMTSNVASVQCDNSDKFPVYKYPKGCGPTLGDGDTLDYYPEDT 300  
Db 241 ENVVWVKQGVVDVKAMTSNVASVQCDNSDKFPVYKYPKGCGPTLGDGDTLDYYPEDT 300  
  
Qy 301 EGDVTRATRTVVKFPTKARPT-----CPP--CPAPEALGA 334  
Db 301 EGDVTRATRTVVKFPTKA--HTPWGLFYFSTASLDEPRGPTIKPCPPCKPAPNLGG 358  
  
Qy 335 PSVLFPPPKDNTLMISRTPTVTCVVDVSHEDDEVKFNWTVGVVHNAKTKPREQYN 394  
Db 359 PSVFIFPKIKDVLMSISPIVTCVVDVSDDDPDVQISWVFNVNVHTAQTOHREDYN 418  
  
Qy 395 STYVWVSLTVLHODLNGKVKCKVSNKALPVPIEKTISKAGQPEQVYTLPPSREE 454  
Db 419 STLKWSALPQHQDMSGKPKCKVNNKDLPAPIERTISKPGSVRAPQVYVLPPEEE 478  
  
Qy 455 MTKQVSLTCLVKGFYPSDIAVEWESNQPPENYKTTTPPVLDSGSPFLYSKLTVDKSR 514  
Db 479 MTKQVTLTCVTDMPEDIYVETWNNKGTLYNKTEPVLDSGSPFLYSKLTVDKSR 538  
  
Qy 515 QQGVFSCSVMEALHNNHYTOKSLSPCK 544  
Db 539 VERNYSYSCSVVHEGLNHHHTKSFSTRTEGK 568

RESULT 9  
AAE12135  
ID AAE12135 standard; Protein; 626 AA.  
XX  
AC AAE12135;  
XX  
XX  
DT 03-JAN-2002 (first entry)  
XX  
DE Human glycoprotein Ib (platelet) alpha (GPIbA).  
XX  
KW Human; haplotyping; glycoprotein Ib (platelet) alpha protein; GPIbA;  
KW Bernard-Soulier syndrome; platelet-type von Willebrand disease; HIV;  
KW Alzheimer's disease; human immunodeficiency virus; SNP;  
KW single nucleotide polymorphism; chromosome 17pter-p12.  
XX  
OS Homo sapiens.  
XX

FF  
FT Misc-difference 72  
FT /note= "Arg at this position is replaced with His  
FT due to single nucleotide polymorphism (SNP)"  
FT  
FT Misc-difference 86  
FT /note= "Leu at this position is replaced with Phe  
FT due to single nucleotide polymorphism (SNP)"  
FT  
FT Misc-difference 161  
FT /note= "Thr at this position is replaced with Met  
FT due to single nucleotide polymorphism (SNP)"  
FT  
FT Misc-difference 592  
FT /note= "Arg at this position is replaced with His  
FT due to single nucleotide polymorphism (SNP)"  
FT  
FT Misc-difference 624  
FT /note= "His at this position is replaced with Arg  
FT due to single nucleotide polymorphism (SNP)"  
FT  
XX  
XX  
PN W0200175065-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 03-APR-2001; 2001WO-US10671.  
XX  
XX 03-APR-2000; 2000US-194341P.  
XX  
XX (GENA-) GENAISSANCE PHARM INC.  
XX  
XX Bentivegna SC, Choi JY, Kliem SE, Koshy B, Parks KE;  
XX  
XX WPI; 2001-626427/72.  
DR N-PSDB; AAD20681.  
DR  
XX  
XX New haplotypes of the glycoprotein Ib platelet alpha polypeptide gene  
PT are useful for diagnosis and drug discovery for treating Bernard  
PT Soulier syndrome, platelet-type von Willebrand disease, HIV and  
PT Alzheimer's disease  
PT  
XX  
XX Claim 29; Fig 3; 66pp; English.  
XX  
XX The invention relates to methods for haplotyping glycoprotein Ib  
XX (platelet) alpha polypeptide (GPIbA) gene of an individual. The  
XX method involves determining if the individual has one of the GPIbA  
XX haplotypes or haplotype pairs. The methods of the invention are  
XX useful for disease diagnosis and in the discovery and development  
XX of drugs for treating diseases associated with GPIbA activity e.g.  
XX Bernard-Soulier syndrome, platelet-type von Willebrand disease, HIV  
XX and Alzheimer's disease. The present sequence is human GPIbA protein.  
XX GPIbA gene is located on chromosome 17pter-p12.  
XX  
SQ Sequence 626 AA;  
  
Query Match 58.3%; Score 1693; DB 22; Length 626;  
Best Local Similarity 87.4%; Pred. No. 3.6e-122;  
Matches 333; Conservative 4; Mismatches 18; Indels 26; Gaps 5;  
  
Qy 1 MPLLLLLLLLPSPHPHPICEVSKVASHLEVNCCKRNLTPDLPKDTTILHLENLY 60  
Db 1 MPLLLLLLLLPSPHPHPICEVSKVASHLEVNCCKRNLTPDLPKDTTILHLENLY 60  
  
Qy 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQQLSLPLGQTLPALTV 120  
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQQLSLPLGQTLPALTV 120  
  
Qy 121 LDVSFNLRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKSLANNLTLP 180  
Db 121 LDVSFNLRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKSLANNLTLP 180  
  
Qy 181 AGLLNGLENDTLLOENSLYTIKPGFSGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240  
Db 181 AGLLNGLENDTLLOENSLYTIKPGFSGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240  
  
Qy 241 ENVVWVKQGVVDVKAMTSNVASVQCDNSDKFPVYKYPKGCGPTLGDGDTLDYYPEDT 300  
Db 241 ENVVWVKQGVVDVKAMTSNVASVQCDNSDKFPVYKYPKGCGPTLGDGDTLDYYPEDT 300

```

Db      241 ENVYVMKQGVQDVKAMTSNVAQVQNSDKFPVYKPGKGCPTLGDGEGDLDLYYPEDT 300
AA89436
QY      301 EGDVRASTRVVKPTKARHTCP-----PCPA---PEALGAPSVLFPPK 343
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      301 EGDVRASTRVVKPTKARHTCP-----PCPA---PEALGAPSVLFPPK 343
QY      344 --PKDTL-----MISRTPEVT 357
        || || || || || || || || || || || || || || || || || || || || ||
Db      359 WTPNFTLHMESITFSKTPKST 379

RESULT 10
ID      AAR89436 standard; Protein; 610 AA.
XX
AC      AAR89436;
XX
DT      25-MAR-2003 (updated)
DT      02-SEP-1996 (first entry)
XX
DE      Mutated platelet glycoprotein-Ib-alpha GPIb protein sequence.
XX
KW      Platelet glycoprotein-Ib-alpha; GPIb; mutagenesis; point mutation;
KW      von Willebrand factor; blood disorder; platelet disorder;
KW      protein engineering; Bernard-Soulier disease.
XX
OS      Homo sapiens.
XX
FH      Key Location/Qualifiers
FT      Misc-difference 57
FT      /note= "substitution from Phe in wild-typ
FT      GPIb-alpha"
FT      Region 36..200
FT      /note= "Leu rich area"
FT      Region 200..220
FT      /note= "flanking region to Leu rich area"
FT      Region 220..310
FT      /note= "hinge region"
FT      Region 310..420
FT      /note= "Ser/Thr rich area"
XX
PN      US5492809-A.
XX
PD      20-FEB-1996.
XX
PF      09-SEP-1993; 93US-0119262.
XX
PR      15-JAN-1992; 92US-0821717.
PR      07-OCT-1991; 91US-0770968.
PR      09-SEP-1993; 93US-0119262.
XX
PA      (UYNV ) UNIV NEW YORK STATE RES FOUND.
XX
PI      Cunningham D, Finch CN, Lyle VA, Miller JL;
XX
DR      W21; 1996-128585/13.
XX
PT      DNA encoding platelet glyco:protein Ib alpha mutant Phe57 -
PT      introduced into platelets to reduce aggregation and reactivity with
PT      von Willebrand factor, also probe for diagnosis of Bernard-Soulier
PT      disease
XX
PS      Disclosure; Column 21-26; 20pp; English.
XX
CC      A substitution of T for C at position 259 in the DNA sequence of
CC      GPIb-alpha leads to the replacement of Phe for Leu at
CC      residue 57 of the mature GPIb-alpha molecule. This mutated GPIb-
CC      alpha protein is less reactive with von Willebrand factor, a
CC      characteristic of the autosomal recessive bleeding disorder,
CC      Bernard-Soulier disease.
CC      (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ      Sequence, 610 AA;

```

```

Query Match 55.0%; Score 1599; DB 17; Length 610;
Best Local Similarity 86.3%; Pred. No. 6.4e-115;
Matches 315; Conservative 4; Mismatches 20; Indels 26; Gaps 5;

QY      17 HPICEVSKVASHLEVNCDKRNLTAAPPDLPKDTTILHLSENLLYTFSLATLMPYTRLTOL 76
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      1 HPICEVSKVASHLEVNCDKRNLTAAPPDLPKDTTILHLSENLLYTFSLATLMPYTRLTOL 60
QY      77 NLDRCETLKLQVDGTLPLVLGTLDLSHNQLQSLPLGOTLPALTVDVSNRLTSLPLGAL 136
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      61 NLDRCETLKLQVDGTLPLVLGTLDLSHNQLQSLPLGOTLPALTVDVSNRLTSLPLGAL 120
QY      137 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTELPAGLLNGLENLDTLLQ 196
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      121 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTELPAGLLNGLENLDTLLQ 180
QY      197 ENSLYTIKPGFPGSHLLPFAFLHGNPWLNCCELLYFRRLQDQNAENVYVWKQGVKAMT 256
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      181 ENSLYTIKPGFPGSHLLPFAFLHGNPWLNCCELLYFRRLQDQNAENVYVWKQGVKAMT 240
QY      257 SNVASVQCDNSDKFPVYKPGKGCPTLGDGEGDLDLYYPEDTEGDKVRASTRVVKFPT 316
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      241 SNVASVQCDNSDKFPVYKPGKGCPTLGDGEGDLDLYYPEDTEGDKVRASTRVVKFPT 300
QY      317 KARPHTCP-----PCPA---PEALGAPSVLFPPK--PKDTL-----MISR 352
        || || || || || || || || || || || || || || || || || || || || ||
DB      301 KA--HTTPWGLFYSWSTASLDSQMPSSLHTQESTKEQTFFPRWTPNFTLHMESITFSK 358
QY      353 TPBVT 357
        ||: |
DB      359 TPKST 363

RESULT 11
ID      AAW18201 standard; protein; 610 AA.
XX
AC      AAW18201;
XX
DT      25-MAR-2003 (updated)
DT      18-AUG-1997 (first entry)
XX
DE      Platelet glycoprotein Ib alpha naturally-occurring wild-type.
XX
KW      GPIb alpha; recombinant polypeptide; antithrombotic;
KW      platelet adhesion; platelet aggregation; thrombus formation;
KW      platelet-type von Willebrand disease.
XX
OS      Homo sapiens.
XX
FH      Key Location/Qualifiers
FT      Region 36..200
FT      /label= Leucine-rich
FT      Region 220..310
FT      /label= Hinge
FT      Region 228..238
FT      /note= "A major binding site for von Willebrand factor"
FT      /note= "Preferred sites for mutations that result
FT      in a mutant polypeptide having more reactivity
FT      with von Willebrand factor"
XX
PS      Misc-difference 233
FT      /note= "Preferred mutation; substitution of Val for Gly"
FT      Region 310..420
FT      /label= Serine/threonine-rich
XX
XX      US5624817-A.
XX
PD      29-APR-1997.
XX
PF      28-APR-1994; 94US-0234265.
XX
PR      07-OCT-1991; 91US-0770968.

```

```

PR 28-APR-1994; 94US-0234265.
XX (UTNY ) UNIV NEW YORK STATE RES FOUND.
XX CUNNINGHAM D, Finch CN, Lyle VA, Miller JL, Pincus MR;
XX WPI; 1997-258215/23.
XX DNA encoding mutant platelet glyco:protein Ib alpha - for production
XX of recombinant polypeptide(s) useful as antithrombotic agents, etc.
XX
XX Disclosure: Column 27-32; 20pp; English.
XX
XX The present sequence represents the naturally-occurring wild-type
XX platelet glycoprotein Ib alpha (GPIb alpha). The polypeptide can have a
XX mutation which makes it more reactive with von Willebrand factor (vWF)
XX than the wild-type. Preferably the mutation occurs in the hinge region,
XX within residues 228 to 238. The preferred mutation is a substitution of
XX Val for the wild-type Gly at position 233. The mutation alters the
XX three-dimensional structure of the mutant polypeptide from a beta-bend
XX conformation to an alpha helix conformation, and also creates an
XX amphipathic region within the polypeptide. The DNA encoding mutant
XX platelet GPIb alpha polypeptides (where the mutation is between the
XX residues 228 to 238, but not at position 233), vectors and cells are
XX useful for production of recombinant mutant GPIb alpha polypeptides,
XX which can be used to inhibit platelet adhesion and aggregation (e.g. in
XX antithrombotic medicaments or on medical devices or vascular
XX prostheses), or can be labelled and used as imaging agents, or can be
XX coupled to a thrombolytic agent for targeting to sites of thrombus
XX formation. The oligonucleotides can be used as probes for diagnosis of
XX platelet-type von Willebrand disease.
XX (Updated on 25-MAR-2003 to correct PF field.)
XX
XX Sequence 610 AA;
SQ
Query Match 55.0%; Score 1599; DB 18; Length 610;
Best Local Similarity 86.3%; Pred. No. 6.4e-115;
Matches 315; Conservative 4; Mismatches 20; Indels 26; Gaps 5;
QY 17 HPICEVSKVASHLEVNCNKRNTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQL 76
DB 1 HPICEVSKVASHLEVNCNKRNTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQL 60
QY 77 NLDRCETLKLQVDTGLPVLTGLDLSHNOQLSLPLGQTLPALTVLDVSFNRLTSLPLGAL 136
DB 61 NLDRCETLKLQVDTGLPVLTGLDLSHNOQLSLPLGQTLPALTVLDVSFNRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196
DB 121 RGLGELQELYLKVNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180
QY 197 ENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEILYFRRLQDQNAENVYWKQGVVKAMT 256
DB 181 ENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEILYFRRLQDQNAENVYWKQGVVKAMT 240
QY 257 SNVASVQCDNSDKFPVYKPGKCGPTLGDGDTLDYDYYPEEDTEGDKVTRATRVVKFPT 316
DB 241 SNVASVQCDNSDKFPVYKPGKCGPTLGDGDTLDYDYYPEEDTEGDKVTRATRVVKFPT 300
QY 317 KARPHTCP-----PCPA---PEALGAPSVFLFPPK--PKDTL-----MISR 352
DB 301 KA--HTTPWGLFYFSWSTASLDQSPSSLHPTQSTKEQITTFPRPWTNPTFLHVESITFSK 358
QY 353 TPEVT 357
DB 359 TPKST 363

```

RESULT 12  
AAR51116  
ID AAR51116 standard; protein; 610 AA.  
XX  
AC AAR51116;

```

XX 25-MAR-2003 (updated)
XX 23-SEP-1994 (first entry)
XX Platelet glycoprotein Ib alpha.
XX
XX Platelet; glycoprotein; imaging; thrombolytic agent;
XX tissue plasminogen activator; tPA; pro-urokinase; urokinase;
XX streptokinase; Bernard-Soulier disease; thrombus; aggregation;
XX anisoylated plasminogen-streptokinase activator complex; adhesion;
XX inhibition.
XX
XX Homo sapiens.
XX
XX US5298239-A.
XX
XX 29-MAR-1994.
XX
XX 15-JAN-1992; 92US-0821717.
XX
XX 07-OCT-1991; 91US-0770968.
XX
XX 15-JAN-1992; 92US-0821717.
XX (UTNY ) UNIV NEW YORK STATE RES FOUND.
XX CUNNINGHAM D, Finch CN, Lyle VA, Miller JL;
XX WPI; 1994-100287/12.
XX
XX Platelet glycoprotein Ib alpha with an amino acid substn at
XX position 57 - has reduced reactivity with Von Willebrand factor,
XX and can be used to inhibit platelet aggregation and inhibition
XX
XX Claim 1; Columns 3-8; 20pp; English.
XX
XX A substitution in platelet glycoprotein Ib alpha (Leucine 57 to
XX Phenylalanine) underlies a form of Bernard-Soulier disease. The
XX mutated glycoprotein can be used in compositions to inhibit
XX platelet aggregation/adhesion. The glycoprotein may be labelled and
XX used as an imaging agent and may also be bound to a thrombolytic
XX agent, preferably tissue plasminogen activator (tPA),
XX (pro) urokinase, streptokinase, anisoylated plasminogen-streptokinase
XX activator complex, tPA analogues or a protease, allowing localisation
XX of the thrombolytic agent to a thrombus.
XX (Updated on 25-MAR-2003 to correct PF field.)
XX
XX Sequence 610 AA;
SQ
Query Match 54.9%; Score 1596; DB 15; Length 610;
Best Local Similarity 86.0%; Pred. No. 1.1e-114;
Matches 314; Conservative 5; Mismatches 20; Indels 26; Gaps 5;
QY 17 HPICEVSKVASHLEVNCNKRNTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQL 76
DB 1 HPICEVSKVASHLEVNCNKRNTALPPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQL 60
QY 77 NLDRCETLKLQVDTGLPVLTGLDLSHNOQLSLPLGQTLPALTVLDVSFNRLTSLPLGAL 136
DB 61 NLDRCETLKLQVDTGLPVLTGLDLSHNOQLSLPLGQTLPALTVLDVSFNRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196
DB 121 RGLGELQELYLKVNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180
QY 197 ENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEILYFRRLQDQNAENVYWKQGVVKAMT 256
DB 181 ENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEILYFRRLQDQNAENVYWKQGVVKAMT 240
QY 257 SNVASVQCDNSDKFPVYKPGKCGPTLGDGDTLDYDYYPEEDTEGDKVTRATRVVKFPT 316
DB 241 SNVASVQCDNSDKFPVYKPGKCGPTLGDGDTLDYDYYPEEDTEGDKVTRATRVVKFPT 300
QY 317 KARPHTCP-----PCPA---PEALGAPSVFLFPPK--PKDTL-----MISR 352

```

```

Db      301 KA--HTTPWCLFYSWSTASLDQMPSSLHPTQESTKEQTTPPRWTPNFTLHVESITFSK 350
      || || |
QY      353 TPEVT 357
      ||: |
Db      359 TPKST 363

RESULT 13
AAR5664
ID      AAR5664 standard; protein; 610 AA.
XX
XX      AAR5664;
XX
XX      25-MAR-2003 (updated)
DT      23-SEP-1994 (first entry)
XX
XX      Mutant platelet glycoprotein Ib alpha.
XX
XX      Platelet; glycoprotein; imaging; thrombolytic agent;
XX      tissue plasminogen activator; tPA; pro-urokinase; urokinase;
XX      streptokinase; Bernard-Soulier disease; thrombus; aggregation;
XX      anisoylated plasminogen-streptokinase activator complex; adhesion;
XX      inhibition.
XX
XX      Homo sapiens.
XX
XX      U55298239-A.
PN
XX
XX      29-MAR-1994.
PD
XX
XX      15-JAN-1992; 92US-0821717.
PF
XX
XX      07-OCT-1991; 91US-0770968.
PR
XX      15-JAN-1992; 92US-0821717.
PR
XX
XX      (UYN ) UNIV NEW YORK STATE RES FOUND.
PA
XX
XX      Cunningham D, Finch CN, Lyle VA, Miller JL;
PI
XX      WPI; 1994-100287/12.
DR
XX
XX      Platelet glycoprotein Ib alpha with an amino acid substn at
XX      position 57 - has reduced reactivity with Von Willebrand factor,
XX      and can be used to inhibit platelet aggregation and inhibition
XX
XX      Claim 1; Page ?; 20pp; English.
XX
XX      A substitution in platelet glycoprotein Ib alpha (Leucine 57 to
XX      Phenylalanine) underlies a form of Bernard-Soulier disease. The
XX      mutated glycoprotein can be used in compositions to inhibit
XX      platelet aggregation/adhesion. The glycoprotein may be labelled and
XX      used as an imaging agent and may also be bound to a thrombolytic
XX      agent, preferably tissue plasminogen activator (tPA),
XX      (pro) urokinase, streptokinase, anisoylated plasminogen-streptokinase
XX      activator complex, tPA analogues or a protease, allowing localisation
XX      of the thrombolytic agent to a thrombus.
XX      (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ      Sequence 610 AA;
Query Match 54.8%; Score 1592; DB 15; Length 610;
Best Local Similarity 85.8%; Pred. No. 2.2e-114;
Matches 313; Conservative 5; Mismatches 21; Indels 26; Gaps 5;

Qy      17 HPICEVKVASHLEVNCCKNLTALPPDLKDTTILHLENLLYTSLATMPYTRLTQL 76
      |||||
Db      1 HPICEVKVASHLEVNCCKNLTALPPDLKDTTILHLENLLYTSLATMPYTRFTQL 60
      |||||

Qy      77 NLDRCETLKQVDTGLTVLGTDLDSHNQSLPLGLGTLFALTVDSFNRLTSLPLGAL 136
      |||||
Db      61 NLDRCETLKQVDTGLTVLGTDLDSHNQSLPLGLGTLFALTVDSFNRLTSLPLGAL 120
      |||||

```

```

QY      137 RGLGELQELKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196
      |||||
Db      121 RGLGELQELYKYNELKTLPPGLLTPTPKLEKLSLANNLTLPVGLLNGLENLDTLLQ 180
      |||||
QY      197 ENSLYTIPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRWLQDNAENYVYWKQGVVKAMT 256
      |||||
Db      181 ENSLYTIPKGFSGSHLLPFAFLHGNPWLNCNCEILYFRRWLEDNAENYVYWKQGVVKAMT 240
      |||||
QY      257 SNVASQCDNSDKFPYKYKGCPTLGDGDDTLDYDYPEEDTEGDKVTRATTVVKFPT 316
      |||||
Db      241 SNVASQCDNSDKFPYKYKGCPTLGDGDDTLDYDYPEEDTEGDKVTRATTVVKFPT 300
      |||||
QY      317 KARPHTCP-----PCPA---PEALGAPSVFLFPPK--PKDTL-----MISR 352
      || || |
Db      301 KA--HTTPWGLFYSWSTASLDQMPSSLHPTQESTKEQTTPPRWTPNFTLHVESITFSK 358
      || || |
QY      353 TPEVT 357
      ||: |
Db      359 TPKST 363

RESULT 14
ABB78241
ID      ABB78241 standard; Protein; 302 AA.
XX
XX      ABB78241;
XX
XX      25-NOV-2002 (first entry)
DT
XX
XX      Amino acid sequence of glycoprotein IB-alpha polypeptide GPIb302/2A.
DE
XX
XX      Glycoprotein IB-alpha; GPIb; immunoglobulin; Ig; platelet adherence;
KW      leukocyte; platelet activation; ischaemic heart disease;
KW      acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;
KW      arterial thrombosis; angina; vascular condition; vascular inflammation;
KW      thrombosis; angioplasty; restenosis.
XX
XX      Unidentified.
OS
XX
XX      WO200263003-A2.
PN
XX
XX      15-AUG-2002.
PD
XX
XX      06-FEB-2002; 2002WO-US03549.
PF
XX
XX      06-FEB-2001; 2001US-266838P.
PR
XX
XX      (GENY ) GENETICS INST LLC.
PA
XX
XX      Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;
PI
XX      WPI; 2002-657537/70.
DR
XX
XX      New glycoprotein Ib alpha fusion polypeptides, useful for treating a
XX      PT disorder associated with platelet activation e.g. ischaemic heart
XX      PT disease, stroke, venous or arterial thrombosis or atherosclerosis -
XX      PS Disclosure; Page 10; 45pp; English.
XX
XX      ABB78240-45 represent glycoprotein IB-alpha (GPIb) polypeptides, which
XX      are used to produce fusion proteins with an immunoglobulin (Ig)
XX      polypeptide. The fusion proteins inhibit the adherence of platelets to
XX      CC leukocytes. The fusion polypeptides are useful for treating a disorder
XX      CC associated with platelet activation e.g. ischaemic heart disease, acute
XX      CC myocardial infarction, stroke, venous thrombosis, atherosclerosis,
XX      CC arterial thrombosis or unstable angina. They can also be used to treat
XX      CC vascular conditions associated with vascular inflammation, thrombosis,
XX      CC and angioplasty-related restenosis.
XX
SQ      Sequence 302 AA;
Query Match 54.6%; Score 1586; DB 23; Length 302;
Best Local Similarity 99.3%; Pred. No. 2.5e-114;

```

Matches 300; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 17 HPICEVSVASHLEVNCNKRNLTAAPPDLPKDPTTILHLSENLLYTFSLATLMPYTRLTOL 76  
DB 1 HPICEVSVASHLEVNCNKRNLTAAPPDLPKDPTTILHLSENLLYTFSLATLMPYTRLTOL 60  
QY 77 NLDRCCLTKLQVDGTLPLVLGTLDLSHNQSLPLLGQTLPALTVLDVSFNRLTSLPLGAL 136  
DB 61 NLDRCCLTKLQVDGTLPLVLGTLDLSHNQSLPLLGQTLPALTVLDVSFNRLTSLPLGAL 120  
QY 137 RGLGELQELYLKGNEKLTLPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196  
DB 121 RGLGELQELYLKGNEKLTLPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180  
QY 197 ENSLYTIPKGFSGSHLLPFAFLHGNPMLCNCILYFRRLQDQNAENVYVWKQGVVDVKAMT 256  
DB 181 ENSLYTIPKGFSGSHLLPFAFLHGNPMLCNCILYFRRLQDQNAENVYVWKQGVVDVKAMT 240  
QY 257 SNVASVQCDNSDKFPVYKYPGKCPPTLGDEGDTLDYDYPBEDTEGDKVRATRTVVVFPT 316  
DB 241 SNVASVQCDNSDKFPVYKYPGKCPPTLGDEGDTLDYDYPBEDTEGDKVAATATVVKFPT 300  
QY 317 KA 318  
DB 301 KA 302

RESULT 15  
ABB78240  
ID ABB78240 standard; Protein; 301 AA.  
AC ABB78240;  
DT 25-NOV-2002 (first entry)  
XX Amino acid sequence of glycoprotein 1B-alpha polypeptide GPIb302.  
DE Glycoprotein 1B-alpha; Gplb; immunoglobulin; Ig; platelet adherence;  
KW leukocyte; platelet activation; ischaemic heart disease;  
KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
KW thrombosis; angioplasty; restenosis.  
XX  
OS Unidentified.  
XX  
XX WO200263003-A2.  
XX  
XX 15-AUG-2002.  
XX  
XX 06-FEB-2002; 2002WO-US03549.  
XX  
XX 06-FEB-2001; 2001US-266838P.  
XX  
XX (GENY ) GENETICS INST LLC.  
XX  
XX Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
XX WPI; 2002-657537/70.  
XX  
XX New glycoprotein 1b alpha fusion polypeptides, useful for treating a  
PT disorder associated with platelet activation e.g. ischaemic heart  
PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
XX  
XX Disclosure; Page 10; 45pp; English.

XX ABB78240-45 represent glycoprotein 1B-alpha (GPIb) polypeptides, which  
CC are used to produce fusion proteins with an immunoglobulin (Ig)  
CC polypeptide. The fusion proteins inhibit the adherence of platelets to  
CC leukocytes. The fusion polypeptides are useful for treating a disorder  
CC associated with platelet activation e.g. ischaemic heart disease, acute  
CC myocardial infarction, stroke, venous thrombosis, atherosclerosis,  
CC arterial thrombosis or unstable angina. They can also be used to treat  
CC vascular conditions associated with vascular inflammation, thrombosis,

CC and angioplasty-related restenosis.

XX  
SQ Sequence 301 AA;  
Query Match 54.2%; Score 1576.5; DB 23; Length 301;  
Best Local Similarity 99.3%; Pred. No. 1.3e-113;  
Matches 300; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 17 HPICEVSVASHLEVNCNKRNLTAAPPDLPKDPTTILHLSENLLYTFSLATLMPYTRLTOL 76  
DB 1 HPICEVSVASHLEVNCNKRNLTAAPPDLPKDPTTILHLSENLLYTFSLATLMPYTRLTOL 60  
QY 77 NLDRCCLTKLQVDGTLPLVLGTLDLSHNQSLPLLGQTLPALTVLDVSFNRLTSLPLGAL 136  
DB 61 NLDRCCLTKLQVDGTLPLVLGTLDLSHNQSLPLLGQTLPALTVLDVSFNRLTSLPLGAL 120  
QY 137 RGLGELQELYLKGNEKLTLPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196  
DB 121 RGLGELQELYLKGNEKLTLPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180  
QY 197 ENSLYTIPKGFSGSHLLPFAFLHGNPMLCNCILYFRRLQDQNAENVYVWKQGVVDVKAMT 256  
DB 181 ENSLYTIPKGFSGSHLLPFAFLHGNPMLCNCILYFRRLQDQNAENVYVWKQGVVDVKAMT 240  
QY 257 SNVASVQCDNSDKFPVYKYPGKCPPTLGDEGDTLDYDYPBEDTEGDKVRATRTVVVFPT 316  
DB 241 SNVASVQCDNSDKFPVYKYPGKCPPTLGDEGDTLDYDYPBEDTEGDKVRATRT -WKFPPT 299  
QY 317 KA 318  
DB 300 KA 301

Search completed: October 14, 2003, 06:30:23  
Job time : 327.846 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:31:18 ; Search time 197.358 Seconds  
(without alignments)  
444.136 Million cell updates/sec

Title: US-10-068-426-1

Perfect score: 2906

Sequence: 1 MPLLALLLPSLPHPIC.....MHEALNHYTKSLSPGK 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, AA.\*

- 1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/prodata/2/pubpaa/FCR\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/prodata/2/pubpaa/FCRUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/prodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2906	100.0	544	15	US-10-068-426-1 Sequence 1, Appli
2	2894	99.6	544	15	US-10-068-426-2 Sequence 2, Appli
3	2881	99.1	544	15	US-10-068-426-3 Sequence 3, Appli
4	2825.5	97.2	531	15	US-10-068-426-4 Sequence 4, Appli
5	2819.5	97.0	531	15	US-10-068-426-6 Sequence 6, Appli
6	2812.5	96.8	531	15	US-10-068-426-5 Sequence 5, Appli
7	1598	55.0	302	15	US-10-068-426-7 Sequence 7, Appli
8	1586	54.6	302	15	US-10-068-426-8 Sequence 8, Appli
9	1555.5	53.5	301	15	US-10-068-426-9 Sequence 9, Appli
10	1539	53.0	290	15	US-10-068-426-10 Sequence 10, Appli
11	1533	52.8	290	15	US-10-068-426-12 Sequence 12, Appli
12	1526	52.5	290	15	US-10-068-426-11 Sequence 11, Appli
13	1236	42.5	313	12	US-09-825-580-4 Sequence 4, Appli
14	1236	42.5	313	15	US-10-211-786-4 Sequence 4, Appli
15	1232.5	42.4	388	9	US-09-784-623-16 Sequence 16, Appli

16	1232	42.4	580	11	US-09-972-268-36 Sequence 36, Appli
17	1230.5	42.3	453	9	US-09-802-077-8 Sequence 8, Appli
18	1230.5	42.3	453	9	US-09-802-096-8 Sequence 8, Appli
19	1230.5	42.3	453	11	US-09-925-179-8 Sequence 8, Appli
20	1229.5	42.3	451	9	US-09-920-171-14 Sequence 14, Appli
21	1229.5	42.3	451	9	US-09-920-171-16 Sequence 16, Appli
22	1229.5	42.3	451	9	US-09-920-171-18 Sequence 18, Appli
23	1229.5	42.3	451	11	US-09-925-179-65 Sequence 65, Appli
24	1229.5	42.3	451	11	US-09-925-179-66 Sequence 66, Appli
25	1229.5	42.3	451	12	US-10-113-996-14 Sequence 14, Appli
26	1229.5	42.3	451	12	US-10-113-996-16 Sequence 16, Appli
27	1229.5	42.3	451	12	US-10-113-996-18 Sequence 18, Appli
28	1229.5	42.3	451	12	US-10-292-869-2 Sequence 2, Appli
29	1229.5	42.3	451	12	US-09-792-938-2 Sequence 2, Appli
30	1225	42.2	694	10	US-09-935-868-18 Sequence 18, Appli
31	1225	42.2	694	12	US-10-282-162-18 Sequence 18, Appli
32	1225	42.2	694	15	US-10-287-035-18 Sequence 32, Appli
33	1225	42.2	793	10	US-09-935-868-32 Sequence 32, Appli
34	1225	42.2	793	12	US-10-282-162-32 Sequence 32, Appli
35	1225	42.2	793	15	US-10-287-035-32 Sequence 32, Appli
36	1224.5	42.1	977	12	US-10-357-653-1 Sequence 1, Appli
37	1224	42.1	401	9	US-09-859-361-9 Sequence 9, Appli
38	1223.5	42.1	451	11	US-09-925-179-68 Sequence 68, Appli
39	1223	42.1	382	15	US-10-207-655-307 Sequence 307, App
40	1223	42.1	680	7	US-08-469-583A-15 Sequence 15, Appli
41	1222	42.1	470	15	US-10-264-634-33 Sequence 33, Appli
42	1221.5	42.0	449	16	US-10-323-268-23 Sequence 23, Appli
43	1221	42.0	450	10	US-09-996-288-238 Sequence 238, App
44	1221	42.0	450	10	US-09-996-288-242 Sequence 242, App
45	1221	42.0	450	10	US-09-996-288-244 Sequence 244, App

ALIGNMENTS

RESULT 1

- US-10-068-426-1
- Sequence 1, Application US/10068426
- Publication No. US20030091576A1
- GENERAL INFORMATION:
- APPLICANT: Shaw, Gray D.
- APPLICANT: Sako, Dianne S.
- APPLICANT: Kumar, Ravindra
- APPLICANT: Sullivan, Francis
- APPLICANT: McDonagh, Tom
- TITLE OF INVENTION: Platellet Glycoprotein IB Alpha Fusion Polypeptides and
- TITLE OF INVENTION: Methods of Use Thereof
- FILE REFERENCE: 22058-503
- CURRENT APPLICATION NUMBER: US/10/068,426
- CURRENT FILING DATE: 2002-02-06
- PRIOR APPLICATION NUMBER: 60/266,838
- PRIOR FILING DATE: 2001-02-06
- NUMBER OF SEQ ID NOS: 20
- SOFTWARE: PatentIn Ver. 2.1
- SEQ ID NO 1
- LENGTH: 544
- TYPE: PRT
- ORGANISM: Homo sapiens
- FEATURE:
- NAME/KEY: DOMAIN
- LOCATION: (1)..(544)
- OTHER INFORMATION: GP1b302-Ig

Query Match 100.0%; Score 2906; DB 15; Length 544;

Best Local Similarity 100.0%; Pred. No. 2.2e-224; Mismatches 0; Indels 0; Gaps 0; Matches 544; Conservative 0;

QY 1 MPLLLLLLPSLPHPICEVSKVASHLEVNCCKRNLTPDLPKDTTILHLENLLY 60

Db 1 MPLLLLLLPSLPHPICEVSKVASHLEVNCCKRNLTPDLPKDTTILHLENLLY 60

QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDGTLFVLGTLDLSHNLQSLPLLGOTLPALTU 120

Db 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLVLGTDLSHNQSLPLGQTLPALTV 120  
Qy 121 LDVSFNRSLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180  
Db 121 LDVSFNRSLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180  
Qy 181 AGLLNGLENLDTLLQENS:YTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240  
Db 181 AGLLNGLENLDTLLQENS:YTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240  
Qy 241 ENVYVMKQGVVDVKAMT:SNVASVQCDNSDKFPVYKPKGCPTLGDGSDTLDYDYPEEDT 300  
Db 241 ENVYVMKQGVVDVKAMT:SNVASVQCDNSDKFPVYKPKGCPTLGDGSDTLDYDYPEEDT 300  
Qy 301 EGDKVRATRVVFKPTKARPHTCPPCPAPEALGAPSVFLFPPKPKDTLMISRTPETVCV 360  
Db 301 EGDKVRATRVVFKPTKARPHTCPPCPAPEALGAPSVFLFPPKPKDTLMISRTPETVCV 360  
Qy 361 VDVSHEDEPVKFNWYDGVVHNAKTPREEQNSTYRVVSVLTVLHQDLNKGKEYKCKV 420  
Db 361 VDVSHEDEPVKFNWYDGVVHNAKTPREEQNSTYRVVSVLTVLHQDLNKGKEYKCKV 420  
Qy 421 SNKALPVIPIEKTISKAKQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 480  
Db 421 SNKALPVIPIEKTISKAKQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 480  
Qy 481 NGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL 540  
Db 481 NGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL 540  
Qy 541 SPGK 544  
Db 541 SPGK 544

## RESULT 2

US-10-068-426-2  
; Sequence 2, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaw, Gray D.  
; APPLICANT: Sako, Dianne S.  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Sullivan, Francis  
; APPLICANT: McDonagh, Tom  
; TITLE OF INVENTION: Flatlet Glycoprotein IB Alpha Fusion Polypeptides and  
; FILE REFERENCE: 22058-503  
; CURRENT APPLICATION NUMBER: US/10/068,426  
; PRIOR FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: 60/266,838  
; PRIOR FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 544  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(544)  
; OTHER INFORMATION: GPL302/2A-Ig  
US-10-068-426-2

Query Match 99.6%; Score 2894; DB 15; Length 544;  
Best Local Similarity 99.6%; Pred. No. 2e-223;  
Matches 542; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPPDLPKOTTLILHSENLLY 60  
Db 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPPDLPKOTTLILHSENLLY 60

Qy 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLVLGTDLSHNQSLPLGQTLPALTV 120  
Db 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLVLGTDLSHNQSLPLGQTLPALTV 120  
Qy 121 LDVSFNRSLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180  
Db 121 LDVSFNRSLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180  
Qy 181 AGLLNGLENLDTLLQENS:YTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240  
Db 181 AGLLNGLENLDTLLQENS:YTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240  
Qy 241 ENVYVMKQGVVDVKAMT:SNVASVQCDNSDKFPVYKPKGCPTLGDGSDTLDYDYPEEDT 300  
Db 241 ENVYVMKQGVVDVKAMT:SNVASVQCDNSDKFPVYKPKGCPTLGDGSDTLDYDYPEEDT 300  
Qy 301 EGDKVRATRVVFKPTKARPHTCPPCPAPEALGAPSVFLFPPKPKDTLMISRTPETVCV 360  
Db 301 EGDKVAATATVVKPTKARPHTCPPCPAPEALGAPSVFLFPPKPKDTLMISRTPETVCV 360  
Qy 361 VDVSHEDEPVKFNWYDGVVHNAKTPREEQNSTYRVVSVLTVLHQDLNKGKEYKCKV 420  
Db 361 VDVSHEDEPVKFNWYDGVVHNAKTPREEQNSTYRVVSVLTVLHQDLNKGKEYKCKV 420  
Qy 421 SNKALPVIPIEKTISKAKQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 480  
Db 421 SNKALPVIPIEKTISKAKQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWES 480  
Qy 481 NGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL 540  
Db 481 NGQPENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL 540  
Qy 541 SPGK 544  
Db 541 SPGK 544

## RESULT 3

US-10-068-426-3  
; Sequence 3, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaw, Gray D.  
; APPLICANT: Sako, Dianne S.  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Sullivan, Francis  
; APPLICANT: McDonagh, Tom  
; TITLE OF INVENTION: Flatlet Glycoprotein IB Alpha Fusion Polypeptides and  
; FILE REFERENCE: 22058-503  
; CURRENT APPLICATION NUMBER: US/10/068,426  
; PRIOR FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: 60/266,838  
; PRIOR FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 544  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(544)  
; OTHER INFORMATION: GPL302/4X-Ig  
US-10-068-426-3

Query Match 99.1%; Score 2881; DB 15; Length 544;  
Best Local Similarity 99.3%; Pred. No. 2.2e-222;  
Matches 540; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPPDLPKOTTLILHSENLLY 60  
Db 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCCKRNLTALPPDLPKOTTLILHSENLLY 60



QY 61 TFSLATLMEYTRLTQNLDRCELTKLOVDGTLVLGTLTDLSDHNOQSLPLGQTLPALTV 120  
DB 61 TFSLATLMEYTRLTQNLDRCELTKLOVDGTLVLGTLTDLSDHNOQSLPLGQTLPALTV 120  
QY 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180  
DB 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180  
QY 181 AGLLNGLENLDTLLLOENSLYTIPIKGFSGSHLLPFAFLHGNPMLCNCETILYFRRLQDNA 240  
DB 181 AGLLNGLENLDTLLLOENSLYTIPIKGFSGSHLLPFAFLHGNPMLCNCETILYFRRLQDNA 240  
QY 241 ENVYVWKQGVVDKAMTSNVASVQCDNSDKFPVYKYPGKGCPTLGDGDTLDLYDYYPEEDT 300  
DB 241 ENVYVWKQGVVDKAMTSNVASVQCDNSDKFPVYKYPGKGCPTLGDGDTLDLYDYYPEEDT 300  
QY 301 EGDKVAATATVVKFPTKARPHCTPCPCPAPEALGAPSVFLFPPKPKDTLMISRTPEVTCV 360  
DB 301 EGDKVAATATVVKFPTKARPHCTPCPCPAPEALGAPSVFLFPPKPKDTLMISRTPEVTCV 360  
QY 361 VDVSHPEDPEVKFNWYVDGVEVHNKTKPREEOYNSYRVSUVTVLHODWLNKKEYCKV 420  
DB 361 VDVSHPEDPEVKFNWYVDGVEVHNKTKPREEOYNSYRVSUVTVLHODWLNKKEYCKV 420  
QY 421 SNKALPVPPIEKTISKAKGQPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVEMES 480  
DB 421 SNKALPVPPIEKTISKAKGQPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVEMES 480  
QY 481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSL 540  
DB 481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSL 540  
QY 541 SPKG 544  
DB 541 SPKG 544

RESULT 4  
US-10-068-426-4  
; Sequence 4, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaw, Gray D.  
; APPLICANT: Sako, Dianne S.  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Sullivan, Francis  
; APPLICANT: McDonagh, Tom  
; TITLE OF INVENTION: 'Platlet Glycoprotein IB Alpha Fusion Polypeptides and  
; TITLE OF INVENTION: 'Methods of Use Thereof  
; FILE REFERENCE: 22058-503  
; CURRENT APPLICATION NUMBER: US/10/068,426  
; CURRENT FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: 60/266,838  
; PRIOR FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(531)  
; OTHER INFORMATION: GPlb290-1g  
US-10-068-426-4  
Query Match 97.2%; Score 2825.5; DB 15; Length 531;  
Best Local Similarity 97.6%; Pred. No. 5.9e-218;  
Matches 531; Conservative 0; Mismatches 0; Indels 13; Gaps 1;  
QY 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCCKRNLTAIPDLPKDTTLHLSENLLY 60  
|||||

DB 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCCKRNLTAIPDLPKDTTLHLSENLLY 60  
QY 61 TFSLATLMEYTRLTQNLDRCELTKLOVDGTLVLGTLTDLSDHNOQSLPLGQTLPALTV 120  
DB 61 TFSLATLMEYTRLTQNLDRCELTKLOVDGTLVLGTLTDLSDHNOQSLPLGQTLPALTV 120  
QY 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180  
DB 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKSLANNLTLP 180  
QY 181 AGLLNGLENLDTLLLOENSLYTIPIKGFSGSHLLPFAFLHGNPMLCNCETILYFRRLQDNA 240  
DB 181 AGLLNGLENLDTLLLOENSLYTIPIKGFSGSHLLPFAFLHGNPMLCNCETILYFRRLQDNA 240  
QY 241 ENVYVWKQGVVDKAMTSNVASVQCDNSDKFPVYKYPGKGCPTLGDGDTLDLYDYYPEEDT 300  
DB 241 ENVYVWKQGVVDKAMTSNVASVQCDNSDKFPVYKYPGKGCPTLGDGDTLDLYDYYPEEDT 300  
QY 301 EGDKVAATATVVKFPTKARPHCTPCPCPAPEALGAPSVFLFPPKPKDTLMISRTPEVTCV 360  
DB 301 EGDKVAATATVVKFPTKARPHCTPCPCPAPEALGAPSVFLFPPKPKDTLMISRTPEVTCV 360  
QY 361 VDVSHPEDPEVKFNWYVDGVEVHNKTKPREEOYNSYRVSUVTVLHODWLNKKEYCKV 420  
DB 361 VDVSHPEDPEVKFNWYVDGVEVHNKTKPREEOYNSYRVSUVTVLHODWLNKKEYCKV 420  
QY 421 SNKALPVPPIEKTISKAKGQPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVEMES 480  
DB 421 SNKALPVPPIEKTISKAKGQPREPQVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVEMES 480  
QY 481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSL 540  
DB 481 NGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSL 540  
QY 541 SPKG 544  
DB 528 SPKG 531

RESULT 5  
US-10-068-426-6  
; Sequence 6, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaw, Gray D.  
; APPLICANT: Sako, Dianne S.  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Sullivan, Francis  
; APPLICANT: McDonagh, Tom  
; TITLE OF INVENTION: 'Platlet Glycoprotein IB Alpha Fusion Polypeptides and  
; TITLE OF INVENTION: 'Methods of Use Thereof  
; FILE REFERENCE: 22058-503  
; CURRENT APPLICATION NUMBER: US/10/068,426  
; CURRENT FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: 60/266,838  
; PRIOR FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(531)  
; OTHER INFORMATION: GPlb290/1A-1g  
US-10-068-426-6  
Query Match 97.0%; Score 2819.5; DB 15; Length 531;  
Best Local Similarity 97.4%; Pred. No. 1.8e-217;  
Matches 530; Conservative 0; Mismatches 1; Indels 13; Gaps 1;  
QY 1 MPLLALLLPSPLPHPICEVSKVASHLEVNCCKRNLTAIPDLPKDTTLHLSENLLY 60  
|||||

```
Db 1 MLLLLLLLLLPSLPHPICEVSKVASHLEVNCNRLTALPPDLPKDTILHSENLLY 60
Qy 61 TFSLATLMPYTRLTQNLDRCELTKLQVGTLPVLGTLDSLHNLQSLPLLGOTLPALT 120
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVGTLPVLGTLDSLHNLQSLPLLGOTLPALT 120
Qy 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKSLANNLTLP 180
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKSLANNLTLP 180
Qy 181 AGLNGLENLDTLLQENSLYTIKGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
Db 181 AGLNGLENLDTLLQENSLYTIKGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
Qy 241 ENVYVWQGVVDVKAAMTSNVASVQCDNSDKFPVYKPGKGTLDGEGDTLDLYYPEEDT 300
Db 241 ENVYVWQGVVDVKAAMTSNVASVQCDNSDKFPVYKPGKGTLDGEGDTLDLYYPEEDT 300
Qy 301 EGDVRAIRTRVVKFPTKARPHTCPPCPAPEALGAPSVFLPPKPKDTLMISRTPEVTCV 360
Db 301 EGDV-----RPHTCPPCPAPEALGAPSVFLPPKPKDTLMISRTPEVTCV 360
Qy 361 VDVSHEDEPVKFNWYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 420
Db 361 VDVSHEDEPVKFNWYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 420
Qy 421 SNKALPVPDIETISKAKGQPREPOVYTLPPSRREMTKNQVSLTCLVKGFYPSDIAVEWES 480
Db 408 SNKALPVPDIETISKAKGQPREPOVYTLPPSRREMTKNQVSLTCLVKGFYPSDIAVEWES 467
Qy 481 NGQPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL 540
Db 468 NGQPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL 527
Qy 541 SPKG 544
Db 528 SPKG 531
```

## RESULT 6

```
US-10-068-426-5
; Sequence 5, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 5
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(531)
; OTHER INFORMATION: GPIb290/2V-Ig
US-10-068-426-5
```

Query Match 96.8%; Score 2812.5; DB 15; Length 531;  
Best Local Similarity 97.2%; Pred. No. 6.5e-217;  
Matches 529; Conservative 1; Mismatches 1; Indels 13; Gaps 1;

```
Qy 1 MLLLLLLLLLPSLPHPICEVSKVASHLEVNCNRLTALPPDLPKDTILHSENLLY 60
Db 1 MLLLLLLLLLPSLPHPICEVSKVASHLEVNCNRLTALPPDLPKDTILHSENLLY 60
Qy 61 TFSLATLMPYTRLTQNLDRCELTKLQVGTLPVLGTLDSLHNLQSLPLLGOTLPALT 120
Db 61 TFSLATLMPYTRLTQNLDRCELTKLQVGTLPVLGTLDSLHNLQSLPLLGOTLPALT 120
Qy 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKSLANNLTLP 180
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKSLANNLTLP 180
Qy 181 AGLNGLENLDTLLQENSLYTIKGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
Db 181 AGLNGLENLDTLLQENSLYTIKGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNA 240
Qy 241 ENVYVWQGVVDVKAAMTSNVASVQCDNSDKFPVYKPGKGTLDGEGDTLDLYYPEEDT 300
Db 241 ENVYVWQGVVDVKAAMTSNVASVQCDNSDKFPVYKPGKGTLDGEGDTLDLYYPEEDT 300
Qy 301 EGDVRAIRTRVVKFPTKARPHTCPPCPAPEALGAPSVFLPPKPKDTLMISRTPEVTCV 360
Db 301 EGDV-----RPHTCPPCPAPEALGAPSVFLPPKPKDTLMISRTPEVTCV 360
Qy 361 VDVSHEDEPVKFNWYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 420
Db 361 VDVSHEDEPVKFNWYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 420
Qy 421 SNKALPVPDIETISKAKGQPREPOVYTLPPSRREMTKNQVSLTCLVKGFYPSDIAVEWES 480
Db 408 SNKALPVPDIETISKAKGQPREPOVYTLPPSRREMTKNQVSLTCLVKGFYPSDIAVEWES 467
Qy 481 NGQPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL 540
Db 468 NGQPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL 527
Qy 541 SPKG 544
Db 528 SPKG 531
```

## RESULT 7

```
US-10-068-426-7
; Sequence 7, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 7
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(302)
; OTHER INFORMATION: GPIb302
US-10-068-426-7
```

Query Match 55.0%; Score 1598; DB 15; Length 302;  
Best Local Similarity 100.0%; Pred. No. 6.2e-120;  
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 17 HPICEVSKVASHLEWNCCKRNLTPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQL 76
Db 1 HPICEVSKVASHLEWNCCKRNLTPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQL 60
QY 77 NLDRCELTKLQVDTGLPVLGTLDSHNQLSPLIGQTLPALTVLDVSNFRLTSLPLGAL 136
Db 61 NLDRCELTKLQVDTGLPVLGTLDSHNQLSPLIGQTLPALTVLDVSNFRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKLTLPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196
Db 121 RGLGELQELYLKGNELKLTLPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180
QY 197 ENSLYTTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLWLDQNAENVYVMKQGVVDVKAMT 256
Db 191 ENSLYTTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLWLDQNAENVYVMKQGVVDVKAMT 240
QY 257 SNVASVQCDNSDKFPVYKYPGKGCPTLGDGDTLDYDYPEDTEGDKVRASTRVVKFPT 316
Db 241 SNVASVQCDNSDKFPVYKYPGKGCPTLGDGDTLDYDYPEDTEGDKVRASTRVVKFPT 300
QY 317 KA 318
Db 301 KA 302

RESULT 8
US-10-068-426-8
; Sequence 8, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platellet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; PRIOR FILING DATE: 2002-02-06
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 8
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(302)
; OTHER INFORMATION: GP1b302/2A
US-10-068-426-8

Query Match 54.6%; Score 1586; DB 15; Length 302;
Best Local Similarity 99.3%; Pred. No. 5.7e-119;
Matches 300; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 17 HPICEVSKVASHLEWNCCKRNLTPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQL 76
Db 1 HPICEVSKVASHLEWNCCKRNLTPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQL 60
QY 77 NLDRCELTKLQVDTGLPVLGTLDSHNQLSPLIGQTLPALTVLDVSNFRLTSLPLGAL 136
Db 61 NLDRCELTKLQVDTGLPVLGTLDSHNQLSPLIGQTLPALTVLDVSNFRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKLTLPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196
Db 121 RGLGELQELYLKGNELKLTLPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180
QY 197 ENSLYTTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLWLDQNAENVYVMKQGVVDVKAMT 256
Db 191 ENSLYTTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLWLDQNAENVYVMKQGVVDVKAMT 240
```

```
Db 181 ENSLYTTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLWLDQNAENVYVMKQGVVDVKAMT 240
QY 257 SNVASVQCDNSDKFPVYKYPGKGCPTLGDGDTLDYDYPEDTEGDKVRASTRVVKFPT 316
Db 241 SNVASVQCDNSDKFPVYKYPGKGCPTLGDGDTLDYDYPEDTEGDKVRASTRVVKFPT 300
QY 317 KA 318
Db 301 KA 302

RESULT 9
US-10-068-426-9
; Sequence 9, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platellet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; PRIOR FILING DATE: 2002-02-06
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 9
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(301)
; OTHER INFORMATION: GP1b/4X
US-10-068-426-9

Query Match 53.5%; Score 1555.5; DB 15; Length 301;
Best Local Similarity 98.3%; Pred. No. 1.6e-116;
Matches 297; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 17 HPICEVSKVASHLEWNCCKRNLTPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQL 76
Db 1 HPICEVSKVASHLEWNCCKRNLTPDLPKDTTILHLSNLLYTFSLATLMPYTRLTQL 60
QY 77 NLDRCELTKLQVDTGLPVLGTLDSHNQLSPLIGQTLPALTVLDVSNFRLTSLPLGAL 136
Db 61 NLDRCELTKLQVDTGLPVLGTLDSHNQLSPLIGQTLPALTVLDVSNFRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKLTLPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196
Db 121 RGLGELQELYLKGNELKLTLPGLLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180
QY 197 ENSLYTTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLWLDQNAENVYVMKQGVVDVKAMT 256
Db 181 ENSLYTTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLWLDQNAENVYVMKQGVVDVKAMT 240
QY 257 SNVASVQCDNSDKFPVYKYPGKGCPTLGDGDTLDYDYPEDTEGDKVRASTRVVKFPT 316
Db 241 SNVASVQCDNSDKFPVYKYPGKGCPTLGDGDTLDYDYPEDTEGDKVRASTRVVKFPT 299
QY 317 KA 318
Db 300 KA 301

RESULT 10
US-10-068-426-10
; Sequence 10, Application US/10068426
; Publication No. US20030091576A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platiect Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(290)
; OTHER INFORMATION: Gp1b290
; US-10-068-426-10

Query Match          53.0%; Score 1539; DB 15; Length 290;
Best Local Similarity 100.0%; Pred. No. 3.1e-115;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLKPTTILHLSNLLYTFSLATLMPYTRLTOL 76
DB 1 HPICEVSKVASHLEVNCCKRNLTALPPDLKPTTILHLSNLLYTFSLATLMPYTRLTOL 60

QY 77 NLDRCCLTKLQVDTLPVLGTLDSLHNQSLPLGQTLPTLKVLEKLSLANNLTLPAGLLNGLENLDTLLQ 136
DB 61 NLDRCCLTKLQVDTLPVLGTLDSLHNQSLPLGQTLPTLKVLEKLSLANNLTLPAGLLNGLENLDTLLQ 120

QY 137 RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196
DB 121 RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180

QY 197 ENSLYTTPKFGFSGHLLPPAFLLHGNPWLNCCEILYFRFRLQDPAENYVVKQGVKAMT 256
DB 181 ENSLYTTPKFGFSGHLLPPAFLLHGNPWLNCCEILYFRFRLQDPAENYVVKQGVKAMT 240

QY 257 SNVASVQCDNSDKFPVYKYPKGKCPITLGDGDTLDLYDYYPEEDTEGDKVR 306
DB 241 SNVASVQCDNSDKFPVYKYPKGKCPITLGDGDTLDLYDYYPEEDTEGDKVR 290

RESULT 11
US-10-068-426-12
; Sequence 12, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platiect Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(290)
; OTHER INFORMATION: Gp1b290/2V
; US-10-068-426-11

Query Match          52.5%; Score 1526; DB 15; Length 290;
Best Local Similarity 99.3%; Pred. No. 3.4e-114;
Matches 288; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLKPTTILHLSNLLYTFSLATLMPYTRLTOL 76
DB 1 HPICEVSKVASHLEVNCCKRNLTALPPDLKPTTILHLSNLLYTFSLATLMPYTRLTOL 60

QY 77 NLDRCCLTKLQVDTLPVLGTLDSLHNQSLPLGQTLPTLKVLEKLSLANNLTLPAGLLNGLENLDTLLQ 136
DB 61 NLDRCCLTKLQVDTLPVLGTLDSLHNQSLPLGQTLPTLKVLEKLSLANNLTLPAGLLNGLENLDTLLQ 120

QY 137 RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196
DB 121 RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180

; ORGANISM: Homo sapiens
; TYPE: PRT
```

```

; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(290)
; OTHER INFORMATION: Gp1b290/1A
; US-10-068-426-12

Query Match          52.8%; Score 1533; DB 15; Length 290;
Best Local Similarity 99.7%; Pred. No. 9.5e-115;
Matches 289; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLKPTTILHLSNLLYTFSLATLMPYTRLTOL 76
DB 1 HPICEVSKVASHLEVNCCKRNLTALPPDLKPTTILHLSNLLYTFSLATLMPYTRLTOL 60

QY 77 NLDRCCLTKLQVDTLPVLGTLDSLHNQSLPLGQTLPTLKVLEKLSLANNLTLPAGLLNGLENLDTLLQ 136
DB 61 NLDRCCLTKLQVDTLPVLGTLDSLHNQSLPLGQTLPTLKVLEKLSLANNLTLPAGLLNGLENLDTLLQ 120

QY 137 RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196
DB 121 RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180

QY 197 ENSLYTTPKFGFSGHLLPPAFLLHGNPWLNCCEILYFRFRLQDPAENYVVKQGVKAMT 256
DB 181 ENSLYTTPKFGFSGHLLPPAFLLHGNPWLNCCEILYFRFRLQDPAENYVVKQGVKAMT 240

QY 257 SNVASVQCDNSDKFPVYKYPKGKCPITLGDGDTLDLYDYYPEEDTEGDKVR 306
DB 241 SNVASVQCDNSDKFPVYKYPKGKCPITLGDGDTLDLYDYYPEEDTEGDKVR 290

RESULT 12
US-10-068-426-11
; Sequence 11, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platiect Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(290)
; OTHER INFORMATION: Gp1b290/2V
; US-10-068-426-11

Query Match          52.5%; Score 1526; DB 15; Length 290;
Best Local Similarity 99.3%; Pred. No. 3.4e-114;
Matches 288; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLKPTTILHLSNLLYTFSLATLMPYTRLTOL 76
DB 1 HPICEVSKVASHLEVNCCKRNLTALPPDLKPTTILHLSNLLYTFSLATLMPYTRLTOL 60

QY 77 NLDRCCLTKLQVDTLPVLGTLDSLHNQSLPLGQTLPTLKVLEKLSLANNLTLPAGLLNGLENLDTLLQ 136
DB 61 NLDRCCLTKLQVDTLPVLGTLDSLHNQSLPLGQTLPTLKVLEKLSLANNLTLPAGLLNGLENLDTLLQ 120

QY 137 RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 196
DB 121 RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLLNGLENLDTLLQ 180

; ORGANISM: Homo sapiens
; TYPE: PRT
```

Db 121 RGLGELQELYLKGNELKTLPPGLLTTPKLEKLSLANNLTLPAGLLNGLENLDTLLJQ 180  
QY 197 ENSLYTIPIKGFPGSHLLPFAFLHGNPWLNCNCEILYPRRWLQDAENNVYKQGVVKAMT 256  
Db 181 ENSLYTIPIKGFPGSHLLPFAFLHGNPWLNCNCEILYPRRWLQDAENNVYKQGVVKAMT 240  
QY 257 SNVASVQCDNSDKFPYKYPKGCPTLGDGDTLDYDYPREDTEGDKVR 306  
Db 241 SNVASVQCDNSDKFPYKYPKGCPTLGDGDTLDYDYPREDTEGDKVR 290

## RESULT 13

US-09-825-580-4  
; Sequence 4, Application US/09825580  
; Publication No. US20030166521A1  
; GENERAL INFORMATION:  
; APPLICANT: Eppihimer, Michael J.  
; APPLICANT: Schaub, Robert G.  
; APPLICANT: Harris, Alan  
; TITLE OF INVENTION: Inhibition of Thrombosis by Treatment with  
; FILE REFERENCE: GPN-5398  
; CURRENT APPLICATION NUMBER: US/09/825,580  
; PRIOR FILING DATE: 2001-04-02  
; PRIOR APPLICATION NUMBER: 60/193,787  
; PRIOR FILING DATE: 2000-03-31  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 313  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-825-580-4

Query Match 42.5%; Score 1236; DB 12; Length 313;  
Best Local Similarity 88.6%; Pred. No. 6.7e-91;  
Matches 233; Conservative 3; Mismatches 17; Indels 10; Gaps 1;  
QY 292 YDYYPEEDTEGDKVRATRTV-----VKPPTKARHTCPPCPAPEALGAPSVFLFP 341  
Db 51 YDFLPETEPPEMLRNSTDTTPLTGPCTESTTVEPAARHTCPPCPAPEALGAPSVFLFP 110  
QY 342 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNSTYRVVS 401  
Db 111 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNSTYRVVS 170  
QY 402 VLTVLHODWLNGKEYKCKVSNKALPVPPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVS 461  
Db 171 VLTVLHODWLNGKEYKCKVSNKALPVPPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVS 230  
QY 462 LTCLVKGFPSPDIADVWESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRWQQGNVFS 521  
Db 231 LTCLVKGFPSPDIADVWESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRWQQGNVFS 290  
QY 522 CSMVHEALHNHYTKLSLSPGK 544  
Db 291 CSMVHEALHNHYTKLSLSPGK 313

## RESULT 14

US-10-211-786-4  
; Sequence 4, Application US/10211786  
; Publication No. US20030083258A1  
; GENERAL INFORMATION:  
; APPLICANT: Michael J. Eppihimer  
; APPLICANT: Robert G. Schaub  
; APPLICANT: Ronald Tuma  
; TITLE OF INVENTION: MODULATION OF LEUKOCYTE-ENDOTHELIAL INTERACTIONS FOLLOWING ISCHEMIA  
; FILE REFERENCE: B702.0099-00000  
; CURRENT APPLICATION NUMBER: US/10/211,786  
; CURRENT FILING DATE: 2002-08-02  
; PRIOR APPLICATION NUMBER: 60/309,816

; PRIOR FILING DATE: 2001-08-03  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 313  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-211-786-4

Query Match 42.5%; Score 1236; DB 15; Length 313;  
Best Local Similarity 88.6%; Pred. No. 6.7e-91;  
Matches 233; Conservative 3; Mismatches 17; Indels 10; Gaps 1;  
QY 292 YDYYPEEDTEGDKVRATRTV-----VKPPTKARHTCPPCPAPEALGAPSVFLFP 341  
Db 51 YDFLPETEPPEMLRNSTDTTPLTGPCTESTTVEPAARHTCPPCPAPEALGAPSVFLFP 110  
QY 342 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNSTYRVVS 401  
Db 111 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNSTYRVVS 170  
QY 402 VLTVLHODWLNGKEYKCKVSNKALPVPPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVS 461  
Db 171 VLTVLHODWLNGKEYKCKVSNKALPVPPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVS 230  
QY 462 LTCLVKGFPSPDIADVWESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRWQQGNVFS 521  
Db 231 LTCLVKGFPSPDIADVWESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRWQQGNVFS 290  
QY 522 CSMVHEALHNHYTKLSLSPGK 544  
Db 291 CSMVHEALHNHYTKLSLSPGK 313

## RESULT 15

US-09-784-623-16  
; Sequence 16, Application US/09784623  
; Patent No. US20020009454A1  
; GENERAL INFORMATION:  
; APPLICANT: Boone, Thomas C.  
; APPLICANT: Hershenson, Susan  
; APPLICANT: Bevilacqua, Michael P.  
; APPLICANT: Collins, David S.  
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY  
; TITLE OF INVENTION: DISEASES  
; FILE REFERENCE: A-365F  
; CURRENT APPLICATION NUMBER: US/09/784,623  
; CURRENT FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: 09/131,247  
; PRIOR FILING DATE: 1998-08-07  
; PRIOR APPLICATION NUMBER: PCT/US 97/02131  
; PRIOR FILING DATE: 1997-02-10  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 388  
; TYPE: PRT  
; ORGANISM: Human  
US-09-784-623-16

Query Match 42.4%; Score 1232.5; DB 9; Length 388;  
Best Local Similarity 63.1%; Pred. No. 1.7e-90;  
Matches 250; Conservative 21; Mismatches 70; Indels 55; Gaps 6;  
QY 171 LANNLTLPAGLLNG-----LENLDTLLIQNSLYTIPIKGFPGSHLLPFAFLHGNPWL 225  
Db 26 LRNN---QLVAGYLQGNVNLBEEKIDVVPTEPHAL-----FLGIHGRMCL 68  
QY 226 NCEILYPRRWLQDAENNVYKQGVVKAMTSNVASVQCDNSDKFPYKYPKGCPCP---- 281  
Db 69 SCVKSGDETRLQLEAVNI-----TDLSSNRKODKRFATIRSDSGPTTSFESACPCN 122  
QY 282 -----TLGDGDTLDYDYPREDTEGDKVRATRTVVKPPTKARHTCPPCPA 328

Db	123	CTAMEADQPVS	LTWMPDEGV	WTKFYQ	DEEAAEPK	SSDKT-----	HTCPCPA	172
Qy	329	PEALGAPSV	FLFPPKPKD	TLMISRTPE	VTWVVDV	SHEDPEV	KFNWYD	GVVHNAKTP
Db	173	PELIGGPS	VFLFPPKPKD	TLMISRTPE	VTWVVDV	SHEDPEV	KFNWYD	GVVHNAKTP
Qy	389	REEOYNSTY	RVVSVLTVL	HQDWLNG	KEYCKV	SNKALPV	PIEKTISK	AKGQPREPOVYTL
Db	233	REEOYNSTY	RVVSVLTVL	HQDWLNG	KEYCKV	SNKALPV	PIEKTISK	AKGQPREPOVYTL
Qy	449	PPSREEMTK	QVSLTCLV	KGFYPSD	IAVWESN	GQPENNYK	TPPVLD	SDSGSFFLYSKLT
Db	293	PPSRDELTK	QVSLTCLV	KGFYPSD	IAVWESN	GQPENNYK	TPPVLD	SDSGSFFLYSKLT
Qy	509	VDKSRWQ	QGNVFS	CSVMHEAL	HNHYTQ	KSLSL	SPGK	544
Db	353	VDKSRWQ	QGNVFS	CSVMHEAL	HNHYTQ	KSLSL	SPGK	388

Search completed: October 14, 2003, 06:48:21  
Job time : 199.358 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:17:26 ; Search time 22.266 Seconds  
(without alignments)  
1033.731 Million cell updates/sec

Title: US-10-068-426-1

Perfect score: 2906

Sequence: 1 MPLLILLLLPLPHFPC.....MHEALHNNHYTKSLSPGK 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/protdata/2/aa/5A-COMB.pep:\*  
2: /cgn2\_6/protdata/2/aa/5B-COMB.pep:\*  
3: /cgn2\_6/protdata/2/aa/6A-COMB.pep:\*  
4: /cgn2\_6/protdata/2/aa/6B-COMB.pep:\*  
5: /cgn2\_6/protdata/2/aa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/protdata/2/aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1606	55.3	320	1	US-07-613-083B-1
2	1599	55.0	610	1	US-07-821-717B-6
3	1599	55.0	610	1	US-08-119-262B-6
4	1599	55.0	610	1	US-08-135-929A-11
5	1599	55.0	610	1	US-08-234-265A-11
6	1236	42.5	313	3	US-08-713-558F-36
7	1232.5	42.4	388	3	US-09-131-247-16
8	1230.5	42.3	453	3	US-08-466-151-8
9	1230.5	42.3	453	4	US-08-466-163B-8
10	1229.5	42.3	451	2	US-08-887-352B-14
11	1229.5	42.3	451	2	US-08-887-352B-16
12	1229.5	42.3	451	2	US-08-887-352B-18
13	1229.5	42.3	451	3	US-08-466-151-65
14	1229.5	42.3	451	3	US-09-109-207C-14
15	1229.5	42.3	451	3	US-09-109-207C-16
16	1229.5	42.3	451	3	US-09-109-207C-18
17	1229.5	42.3	451	3	US-09-282-505-2
18	1229.5	42.3	451	3	US-09-054-255-2
19	1229.5	42.3	451	3	US-09-296-005-14
20	1229.5	42.3	451	3	US-09-296-005-16
21	1229.5	42.3	451	3	US-09-296-005-18
22	1229.5	42.3	451	4	US-09-282-846-2
23	1229.5	42.3	451	4	US-09-680-145-2
24	1225	42.2	694	4	US-09-313-942-18
25	1225	42.2	793	4	US-09-313-942-32
26	1224.5	42.1	977	4	US-09-590-656-1
27	1224.5	42.1	977	4	US-09-733-764-1

Sequence 15, Appl  
Sequence 23, Appl  
Sequence 23, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 30, Appl  
Sequence 43, Appl  
Sequence 43, Appl  
Sequence 43, Appl  
Sequence 43, Appl  
Sequence 8, Appl  
Sequence 43, Appl  
Sequence 71, Appl  
Sequence 71, Appl  
Sequence 71, Appl  
Sequence 71, Appl  
Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-07-613-083B-1  
; Sequence 1, Application US/07613083B  
; Patent No. 5340727  
; GENERAL INFORMATION:  
; APPLICANT: Ruggeri, Zaverio M.  
; APPLICANT: Ware, Jerry, Inventors  
; APPLICANT: on behalf of Scripps Clinic and Research  
; APPLICANT: Foundation  
; TITLE OF INVENTION: GPIB' Fragments and Recombinant  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scripps Clinic and Research  
; ADDRESS: Foundation  
; STREET: 10666 No. 5340727th Torrey Pines Road  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette-5.25 inch, 1.2 Mb  
; COMPUTER: AST Bravo IBM PC comp. (386SX)  
; OPERATING SYSTEM: MS DOS version 3.2  
; SOFTWARE: WordPerfect 5.1 conv. to ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/613,083B  
; FILING DATE: 19911114  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA: This appl. is a c-i-p of  
; APPLICATION NUMBER: U.S. 07/470,674  
; FILING DATE: 04-Jan-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Barron, Alexis  
; REGISTRATION NUMBER: 22,702  
; REFERENCE/DOCKET NUMBER: P16,569-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 923-4466  
; TELEFAX: (215) 923-2189  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 320  
; TYPE: AMINO ACID  
; STRANDEDNESS: No. 5340727 applicable  
; TOPOLOGY: Linear  
; US-07-613-083B-1

Query Match 55.3%; Score 1606; DB 1; Length 320;  
Best Local Similarity 99.0%; Pred. No. 1.5e-132;  
Matches 305; Conservative 0; Mismatches 1; Indels 2; Gaps 1;



QY 17 HPICEVSKVASHLEVNCNDRNLTALPPDLPKDPTTLHLSENLLYTFSLATLMPYTRLTQL 76  
DB 1 HPICEVSKVASHLEVNCNDRNLTALPPDLPKDPTTLHLSENLLYTFSLATLMPYTRLTQL 60  
QY 77 NLDRCBLTKLQVDTLPVLGTLDSLHNQLOSLPLGQTLPALTVLDVSNRLTSLPLGAL 136  
DB 61 NLDRCBLTKLQVDTLPVLGTLDSLHNQLOSLPLGQTLPALTVLDVSNRLTSLPLGAL 120  
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196  
DB 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180  
QY 197 ENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCNCELTYFRRLQNAENYVWKQGVVKAMT 256  
DB 181 ENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCNCELTYFRRLQNAENYVWKQGVVKAMT 240  
QY 257 SNVASVQCDNSDKFPYKYPKGKPTLGDEGDTLDYYPEDTEGDKVRAIRTVVKFPT 316  
DB 241 SNVASVQCDNSDKFPYKYPKGKPTLGDEGDTLDYYPEDTEGDKVRAIRTVVKFPT 300  
QY 317 KARPHTCP 324  
DB 301 KA--HTTP 306

## RESULT 2

US-07-821-717B-6  
; Sequence 6, Application US/07821717B  
; Patent No. 5298239  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Cunningham, David  
; APPLICANT: Lyle, Vicki A.  
; APPLICANT: Finch, Clara N.  
; TITLE OF INVENTION: MUTATIONS RENDERING PLATELET  
; TITLE OF INVENTION: GLYCOPROTEIN IB ALPHA LESS REACTIVE  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/821,717B  
; FILING DATE: 15-JAN-1992  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Timain, Susan J.  
; REGISTRATION NUMBER: 34,103  
; REFERENCE/DOCKET NUMBER: 20884/21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1636  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PUBLICATION INFORMATION:  
; AUTHORS: Lopez, Jose A.  
; AUTHORS: Chung, Dominic W.  
; AUTHORS: Fujikawa, Kazuo  
; AUTHORS: Hagen, Frederick S.

; AUTHORS: Papayannopoulou, Thalia  
; AUTHORS: Roth, Gerald J.  
; TITLE: Cloning of the alpha chain of human  
; TITLE: Platelet glycoprotein Ib: A transmembrane protein with homology  
; TITLE: to leucine-rich alpha-2-glycoprotein  
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
; VOLUME: 84  
; PAGES: 5615-5619  
; DATE: AUG-1987  
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 610  
; PUBLICATION INFORMATION:  
; AUTHORS: Zimmerman, Theodore S.  
; AUTHORS: Ruggeri, Zaverio M.  
; AUTHORS: Houghten, Richard A.  
; AUTHORS: Vincete, Vincete  
; AUTHORS: Mohri, Hiroshi  
; TITLE: Proteolytic fragments and synthetic  
; TITLE: peptides that block the binding of von Willebrand factor to the  
; TITLE: platelet membrane glycoprotein Ib  
; DOCUMENT NUMBER: EP 0 317 278 A2  
; FILING DATE: 16-NOV-1988  
; PUBLICATION DATE: 24-MAY-1989  
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 293  
; US-07-821-717B-6  
Query Match 55.0%; Score 1599; DB 1; Length 610;  
Best Local Similarity 86.3%; Pred. No. 1.6e-131;  
Matches 315; Conservative 4; Mismatches 20; Indels 26; Gaps 5;  
QY 17 HPICEVSKVASHLEVNCNDRNLTALPPDLPKDPTTLHLSENLLYTFSLATLMPYTRLTQL 76  
DB 1 HPICEVSKVASHLEVNCNDRNLTALPPDLPKDPTTLHLSENLLYTFSLATLMPYTRLTQL 60  
QY 77 NLDRCBLTKLQVDTLPVLGTLDSLHNQLOSLPLGQTLPALTVLDVSNRLTSLPLGAL 136  
DB 61 NLDRCBLTKLQVDTLPVLGTLDSLHNQLOSLPLGQTLPALTVLDVSNRLTSLPLGAL 120  
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196  
DB 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180  
QY 197 ENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCNCELTYFRRLQNAENYVWKQGVVKAMT 256  
DB 181 ENSLYTIPIKGFSGSHLLPFAFLHGNPWLNCNCELTYFRRLQNAENYVWKQGVVKAMT 240  
QY 257 SNVASVQCDNSDKFPYKYPKGKPTLGDEGDTLDYYPEDTEGDKVRAIRTVVKFPT 316  
DB 241 SNVASVQCDNSDKFPYKYPKGKPTLGDEGDTLDYYPEDTEGDKVRAIRTVVKFPT 300  
QY 317 KARPHTCP-----PCPA---PEALGAPSVFLPPPK--PKDTL-----MISR 352  
DB 301 KA--HTTPWGLPYSMSTASLDSQMPSSLHPTQBSTKEQTTFPPRWTNFTLHMBESITFSK 358  
QY 353 TPEVT 357  
DB 359 TPKST 363  
RESULT 3  
US-08-119-262B-6  
; Sequence 6, Application US/08119262B  
; Patent No. 5492809  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Cunningham, David  
; APPLICANT: Lyle, Vicki A.  
; APPLICANT: Finch, Clara N.  
; TITLE OF INVENTION: MUTATIONS RENDERING PLATELET  
; TITLE OF INVENTION: GLYCOPROTEIN IB ALPHA LESS REACTIVE  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051

[illegible]

CITY:	Rochester
STATE:	New York
COUNTRY:	USA
ZIP:	14603
COMPUTER READABLE FORM:	
MEDIUM TYPE:	Floppy disk
OPERATING SYSTEM:	PC-DOS/MS-DOS
SOFTWARE:	PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:	
APPLICATION NUMBER:	US/08/119,262B
FILING DATE:	09-SEP-1993
CLASSIFICATION:	424
PRIOR APPLICATION NUMBER:	US 07/821,717
FILING DATE:	15-JAN-1992
ATTORNEY/AGENT INFORMATION:	
NAME:	Timain, Susan J.
REGISTRATION NUMBER:	34,103
REFERENCE/DOCKET NUMBER:	20884/22
TELEPHONE:	(716) 263-1636
TELEFAX:	(716) 263-1600
INFORMATION FOR SEQ ID NO:	6:
SEQUENCE CHARACTERISTICS:	
LENGTH:	610 amino acids
TYPE:	amino acid
STRANDEDNESS:	
TOPOLOGY:	linear
MOLECULE TYPE:	protein
PUBLICATION INFORMATION:	
AUTHORS:	Lopez, Jose A.
AUTHORS:	Chung, Dominic W.
AUTHORS:	Fujikawa, Kazuo
AUTHORS:	Hagen, Frederick S.
AUTHORS:	Papayannopoulou, Thalia
AUTHORS:	Roth, Gerald J.
TITLE:	Cloning of the alpha chain of human platelet glycoprotein Ib: A transmembrane protein
TITLE:	leucine-rich alpha-2-glycoprotein
JOURNAL:	Proc. Natl. Acad. Sci. U.S.A.
VOLUME:	84
PAGES:	5615-5619
DATE:	AUG-1987
RELEVANT RESIDUES IN SEQ ID NO:	6: FROM 0 TO 610
PUBLICATION INFORMATION:	
AUTHORS:	Zimmerman, Theodore S.
AUTHORS:	Ruggeri, Zaverio M.
AUTHORS:	Houghten, Richard A.
AUTHORS:	Vincete, Vincete
AUTHORS:	Mohri, Hiroshi
TITLE:	Proteolytic fragments and synthetic peptides that block the binding of von Willebrand factor to the membrane glycoprotein Ib
DOCUMENT NUMBER:	EP 0 317 278 A2
FILING DATE:	16-NOV-1988
PUBLICATION DATE:	24-MAY-1989
RELEVANT RESIDUES IN SEQ ID NO:	6: FROM 0 TO 293
US-08-119-262B-6	
Query Match	55.0%; Score 1599; DB 1; Length 610;
Best Local Similarity	86.3%; Pred. No. 1.6e-131;
Matches	315; Conservative 4; Mismatches 20; Indels 26; Gaps 5;
Qy	17 HPICEVKVASHLVNCDKRNLTALPPDLPKDTTILHSENLLYTFSSLATLMPYTRTLQ 76
Dd	1 HPICEVKVASHLVNCDKRNLTALPPDLPKDTTILHSENLLYTFSSLATLMPYTRTLQ 60
Qy	77 NLDRCBLTKLOVDGTLPLVGTLDSHNQLSLGLGQTLPALTVLVDVSFNELTSIPLGAL 136
Dd	61 NLDRCBLTKLOVDGTLPLVGTLDSHNQLSLGLGQTLPALTVLVDVSFNELTSIPLGAL 120
Qy	137 RGLGELQEYLKXNELKTUPPGLLTPTPKLEKLANNNTLTPAGLINGENLDTLLQQ 196

[illegible]

## RESULT 5

```

US-08-234-265A-11
/ Sequence 11, Application US/08234265A
/ Patent No. 5624817
/ GENERAL INFORMATION:
/ APPLICANT: Miller, Jonathan L.
/ APPLICANT: Cunningham, David
/ APPLICANT: Lyle, Vicki A.
/ APPLICANT: Fincus, Matthew R.
/ APPLICANT: Fincus, Matthew R.
/ TITLE OF INVENTION: Mutations in the Gene Encoding the
/ TITLE OF INVENTION: Chain of Platelet Glycoprotein Ib
/ NUMBER OF SEQUENCES: 11
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Nixon, Hargrave, Devans & Doyle
/ STREET: Clinton Square, P.O. Box 1051
/ CITY: Rochester
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 14603
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/234,265A
/ FILING DATE: 28-APR-1994
/ CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Timian, Susan J.
/ REGISTRATION NUMBER: 34,103
/ REFERENCE/DOCKET NUMBER: 20884/24
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (716) 263-1636
/ TELEFAX: (716) 263-1600
/ TELEX: 978450
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 610 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-234-265A-11

```

Query Match	55.0%	Score 1599;	DB 1;	Length 610;
Best Local Similarity	86.3%	Pred. No. 1.6e-131;		
Matches 315; Conservative	4;	Mismatches 20;	Indels 26;	Gaps 5;

Qy	17	HPICEVSVASHLEVNCQKRWLTALPPDLPKOTTILHLSENLLYTSLSLATLMPYTRLTOL	76
Db	1	HPICEVSVASHLEVNCQKRWLTALPPDLPKOTTILHLSENLLYTSLSLATLMPYTRLTOL	60
Qy	77	NLDRCBLTKLQVGDGTLFVLGTLDSLHNOQLSLPLGGTLPALTVDLOVSFNRLTSLPLGAL	136
Db	61	NLDRCBLTKLQVGDGTLFVLGTLDSLHNOQLSLPLGGTLPALTVDLOVSFNRLTSLPLGAL	120
Qy	137	RGELGLOELYLKGNELKTLPPGLTPTPKLEKLSLANNLITELPAGLLNGLENLDTLLQ	196
Db	121	RGELGLOELYLKGNELKTLPPGLTPTPKLEKLSLANNLITELPAGLLNGLENLDTLLQ	180
Qy	197	ENSLYTIKPGFFGSHLLPFAFLHGNPWLNCCEILYFRRWLODNAENVYVWKQGVDKAMT	256
Db	181	ENSLYTIKPGFFGSHLLPFAFLHGNPWLNCCEILYFRRWLODNAENVYVWKQGVDKAMT	240
Qy	257	SNVASVQCDSNDKFPVYKYCKGCGPTLGDGDTLDLYYFEEDETEGDKVRASTRVVKFPT	316
Db	241	SNVASVQCDSNDKFPVYKYCKGCGPTLGDGDTLDLYYFEEDETEGDKVRASTRVVKFPT	300
Qy	317	KARPHTCP-----PCPA---PEALGAPSVFLFPPK--PKOTL-----MISR	352
Db	301	KA--HTTPGLFYSWSTASLDSQMPSSLIHPQESTKEQTTFPRTWPNFTLHMESITFSK	358
Qy	353	TPBVT	357
Db	359	TPKST	363

## RESULT 6

```

US-08-713-556P-36
; Sequence 36, Application US/08713556F
; Patent No. 6277975
; GENERAL INFORMATION:
; APPLICANT: Larsen, Glenn
; APPLICANT: Sako, Dianne
; APPLICANT: Chang, Xiao Jia
; APPLICANT: Veldman, Geertruida M.
; APPLICANT: Cumming, Dale
; APPLICANT: Kumar, Ravindra
; APPLICANT: Shaw, Gray
; TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEGAL AFFAIRS
; STREET: 87 CAMBRIDGEPARK DRIVE
; CITY: CAMBRIDGE
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/713,556P
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/965,662
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/112,608
; FILING DATE: 26-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10168
; FILING DATE: 22-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,398
; FILING DATE: 28-APR-1994
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 08/316,305  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/428,734  
FILING DATE: 25-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BROWN, SCOTT A.  
REGISTRATION NUMBER: 32,724  
REFERENCE/DOCKET NUMBER: GI 5213P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-713-556F-36

Query Match 42.5%; Score 1236; DB 3; Length 313;  
Best Local Similarity 88.6%; Pred. No. 3.6e-100;  
Matches 233; Conservative 3; Mismatches 17; Indels 10; Gaps 1;  
QY 292 YDYYPEEDTECDKVRATRTV-----VKFPTKARPHTCPPCPAPEALGAPSVFLFP 341  
DB 51 YDFLPEPEPMLRNSTDTPLTGPPTSTTVEPARPHTCPPCPAPEALGAPSVFLFP 110  
QY 342 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVS 401  
DB 111 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVS 170  
QY 402 VLTVLHODWLNKGYCKVSKNKPALVPPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVS 461  
DB 171 VLTVLHODWLNKGYCKVSKNKPALVPPIEKTISKAKGQPREPOVYTLPPSREEMTKNQVS 230  
QY 462 LTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFS 521  
DB 231 LTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFS 290  
QY 522 CSMVHEALHNHYTQKSLSLSPCK 544  
DB 291 CSMVHEALHNHYTQKSLSLSPCK 313

RESULT 7  
US-09-131-247-16  
Sequence 16, Application US/09131247  
Patent No. 6294170  
GENERAL INFORMATION:  
APPLICANT: Boone, Thomas C.  
APPLICANT: Herhenson, Susan  
APPLICANT: Revillacqua, Michael P.  
APPLICANT: Collins, David S.  
TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY  
TITLE OF INVENTION: DISEASES  
FILE REFERENCE: A-365P  
CURRENT APPLICATION NUMBER: US/09/131,247  
CURRENT FILING DATE: 1998-08-07  
EARLIER APPLICATION NUMBER: 60/055,185  
EARLIER FILING DATE: 1997-08-08  
EARLIER APPLICATION NUMBER: PCT/US 97/02131  
EARLIER FILING DATE: 1997-02-10  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 16  
LENGTH: 388  
TYPE: PRT  
ORGANISM: Human  
US-09-131-247-16  
Query Match 42.4%; Score 1232.5; DB 3; Length 388;

Best Local Similarity 63.1%; Pred. No. 1e-99;  
Matches 250; Conservative 21; Mismatches 70; Indels 55; Gaps 6;  
QY 171 LANNLTETLPAGLNG-----LENLDTLLOENSLYTIPIKGFSGHLLPFAFLHGNPWL 225  
DB 26 LNNN---QLVAGYLQGNVNLKIDVVPPIEPHAL-----FLGIHGKMKCL 68  
QY 226 NCEILYFRWLNQDAENYVWKQGVDKAMTSNVASVQCDNSDKFPVYKYPKGCPCP----- 281  
DB 69 SCVKSGETRLQLEAVNL-----TDLSENKQDKRFAFIRSDSGPTTSFSAACPGWFL 122  
QY 282 -----TLGDEGDTLDYDYPEEDTEGDKVRATRTVVKFPTKARPHTCPPCPA 328  
DB 123 CTAMEADQPVSLTNMPDEGVVWTKFYFQDEAAEAPKSSDKT-----HTCPCCPA 172  
QY 329 PEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 388  
DB 173 PELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 232  
QY 389 REEQYNSTYRVVSVLTVLHODWLNKGYCKVSKNKPALVPPIEKTISKAKGQPREPOVYTL 448  
DB 233 REEQYNSTYRVVSVLTVLHODWLNKGYCKVSKNKPALVPPIEKTISKAKGQPREPOVYTL 292  
QY 449 PPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLT 508  
DB 293 PPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLT 352  
QY 509 VDKSRWQQGNVFS CSMVHEALHNHYTQKSLSLSPCK 544  
DB 353 VDKSRWQQGNVFS CSMVHEALHNHYTQKSLSLSPCK 388

RESULT 8  
US-08-466-151-8  
Sequence 8, Application US/08466151  
Patent No. 6037453  
GENERAL INFORMATION:  
APPLICANT: Jardieu, Paula M.  
APPLICANT: Presta, Leonard G.  
TITLE OF INVENTION: Immunoglobulin Variants  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESS: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,151  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/466163  
FILING DATE: 06-Jun-1995  
APPLICATION NUMBER: 08/405617  
FILING DATE: 15-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/185899  
FILING DATE: 26-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/879495  
FILING DATE: 07-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/744768  
FILING DATE: 14-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svoboda, Craig G.

```
/ ; REGISTRATION NUMBER: 39,044
/ ; REFERENCE/DOCKET NUMBER: P0718P2C1D1
/ ; TELECOMMUNICATION INFORMATION:
/ ; TELEPHONE: 650/225-1489
/ ; TELEFAX: 650/952-9881
/ ; INFORMATION FOR SEQ ID NO: 8:
/ ; SEQUENCE CHARACTERISTICS:
/ ; LENGTH: 453 amino acids
/ ; TYPE: Amino Acid
/ ; TOPOLOGY: Linear
/ ;
US-08-466-151-8

Query Match 42.3%; Score 1230.5; DB 3; Length 453;
Best Local Similarity 61.7%; Pred. No. 1.9e-99;
Matches 254; Conservative 18; Mismatches 53; Indels 87; Gaps 8;

QY 188 ENLDTLLQLQNSLYTTPKGGFF-----GSHLLPFAFLHGNPWLNCNCEILYFRRWLQDQNAENV 243
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 DSKNTFYLQWNSLRAEDTAVVYCARGSH-----YFGHW-----HF 108

QY 244 YWKQGVVDVKAMTSNVASVQCDNSDKFPV---YKYPGKGCPTLGDGDTCLDYDYPEDT 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 109 AVWQGT---LVTVSSASTKGKGPSVFLPAPSKSTSGGTAALG---CLVKDYFPEPVT 161

QY 301 -----EGDKVRATRTVWKFPFKA----- 318
Db 162 VSNNSGALTSVHTTFAVLQSSGLYSLSVTVTPSSSLGTQYIYCNVNHKPSNTKVDDKV 221

QY 319 -----RPHTCPPCPAPEALGAPSVLFPFKPKDITLMISRTPEVTCVVVDVSHEDPEVKF 372
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 222 EPKSCDKTHTCPPCPAPEALGAPSVLFPFKPKDITLMISRTPEVTCVVVDVSHEDPEVKF 281

QY 373 NMVVDGVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIETK 432
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 282 NMVVDGVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIETK 341

QY 433 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 492
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 342 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 401

QY 493 PVLDSGGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNNHYTKSLSPGK 544
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 402 PVLDSGGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNNHYTKSLSPGK 453

RESULT 9
US-08-466-163B-8
; Sequence 8, Application US/08466163B
; Patent No. 6329509
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; TITLE OF INVENTION: Immunoglobulin Variants
; FILE REFERENCE: P0718P2C1D1
; CURRENT APPLICATION NUMBER: US/08/466,163B
; PRIOR FILING DATE: 1995-06-06
; PRIOR FILING DATE: 1995-03-15
; PRIOR FILING DATE: 1994-01-26
; PRIOR FILING DATE: 1992-05-07
; PRIOR FILING DATE: 1991-08-14
; NUMBER OF SEQ ID NOS: 64
; SEQ ID NO 8
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: humanized mae11, version 1 heavy chain
US-08-466-163B-8

Query Match 42.3%; Score 1229.5; DB 2; Length 451;

/ ; REGISTRATION NUMBER: 39,044
/ ; REFERENCE/DOCKET NUMBER: P0718P2C1D1
/ ; TELECOMMUNICATION INFORMATION:
/ ; TELEPHONE: 650/225-1489
/ ; TELEFAX: 650/952-9881
/ ; INFORMATION FOR SEQ ID NO: 8:
/ ; SEQUENCE CHARACTERISTICS:
/ ; LENGTH: 453 amino acids
/ ; TYPE: Amino Acid
/ ; TOPOLOGY: Linear
/ ;
US-08-466-151-8

Query Match 42.3%; Score 1230.5; DB 4; Length 453;
Best Local Similarity 61.7%; Pred. No. 1.9e-99;
Matches 254; Conservative 18; Mismatches 53; Indels 87; Gaps 8;

QY 188 ENLDTLLQLQNSLYTTPKGGFF-----GSHLLPFAFLHGNPWLNCNCEILYFRRWLQDQNAENV 243
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 DSKNTFYLQWNSLRAEDTAVVYCARGSH-----YFGHW-----HF 108

QY 244 YWKQGVVDVKAMTSNVASVQCDNSDKFPV---YKYPGKGCPTLGDGDTCLDYDYPEDT 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 109 AVWQGT---LVTVSSASTKGKGPSVFLPAPSKSTSGGTAALG---CLVKDYFPEPVT 161

QY 301 -----EGDKVRATRTVWKFPFKA----- 318
Db 162 VSNNSGALTSVHTTFAVLQSSGLYSLSVTVTPSSSLGTQYIYCNVNHKPSNTKVDDKV 221

QY 319 -----RPHTCPPCPAPEALGAPSVLFPFKPKDITLMISRTPEVTCVVVDVSHEDPEVKF 372
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 222 EPKSCDKTHTCPPCPAPEALGAPSVLFPFKPKDITLMISRTPEVTCVVVDVSHEDPEVKF 281

QY 373 NMVVDGVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIETK 432
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 282 NMVVDGVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIETK 341

QY 433 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 492
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 342 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 401

QY 493 PVLDSGGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNNHYTKSLSPGK 544
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 402 PVLDSGGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNNHYTKSLSPGK 453

RESULT 10
US-08-887-352B-14
; Sequence 14, Application US/08887352B
; Patent No. 5994511
; GENERAL INFORMATION:
; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe
; TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/887,352B
; FILING DATE: 03-Jul-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P1123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-887-352B-14
```

Best Local Similarity 61.7%; Pred. No. 2.3e-99;  
Matches 254; Conservative 16; Mismatches 53; Indels 89; Gaps 8;

```

QY 188 ENLDTLLOENSLYTIKGF---GSHLLPFAFLHGNPWLNCCEILYFRRLQDAENV 243
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 74 DSKNTFYLMNSLRAEDTAVVYCARGSH-----YFGHW-----HF 108
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 244 YVMKQGVVDKAMTSNVASQCDNSDKFPV---YKYPGKGCPITLGDGDTLDLYDYPEEDT 300
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 109 AVWGQGTIV-----TVSSASTKGPSVFLPAPSSKSTSGGTAALG-----CLVKDYFPEPVT 159
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 301 -----EGDKVRAIRTVVVFPTKA----- 318
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 160 VSMNSGALTSVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHPKSNTKVDKKV 219
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 319 -----RPHTCPPCPAPALGAPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 372
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 220 EPKSCDKTHTCCPAPPELLGGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 279
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 373 NWYVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 432
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 280 NWYVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 339
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 433 ISKAKGQPREPVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTP 492
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 340 ISKAKGQPREPVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTP 399
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 493 PVLDSGDFLYSKLTVDKSRWQOGNVFSCSVWHEALHNNHYTKSLSPGK 544
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 400 PVLDSGDFLYSKLTVDKSRWQOGNVFSCSVWHEALHNNHYTKSLSPGK 451
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

```

RESULT 11  
US-08-887-352B-16  
; Sequence 16, Application US/08887352B  
; Patent No. 5994511  
; GENERAL INFORMATION:  
; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe  
; TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of  
; TITLE OF INVENTION: Improving Polypeptides  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/887,352B  
; FILING DATE: 03-Jul-1997  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Svoboda, Craig G.  
; REGISTRATION NUMBER: 39,044  
; REFERENCE/DOCKET NUMBER: P1123  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1489  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 451 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-887-352B-16

Query Match 42.3%; Score 1229.5; DB 2; Length 451;  
Best Local Similarity 61.7%; Pred. No. 2.3e-99;  
Matches 254; Conservative 16; Mismatches 53; Indels 89; Gaps 8;

Matches 254; Conservative 16; Mismatches 53; Indels 89; Gaps 8;

```

QY 188 ENLDTLLOENSLYTIKGF---GSHLLPFAFLHGNPWLNCCEILYFRRLQDAENV 243
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 74 DSKNTFYLMNSLRAEDTAVVYCARGSH-----YFGHW-----HF 108
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 244 YVMKQGVVDKAMTSNVASQCDNSDKFPV---YKYPGKGCPITLGDGDTLDLYDYPEEDT 300
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 109 AVWGQGTIV-----TVSSASTKGPSVFLPAPSSKSTSGGTAALG-----CLVKDYFPEPVT 159
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 301 -----EGDKVRAIRTVVVFPTKA----- 318
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 160 VSMNSGALTSVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHPKSNTKVDKKV 219
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 319 -----RPHTCPPCPAPALGAPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 372
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 220 EPKSCDKTHTCCPAPPELLGGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 279
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 373 NWYVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 432
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 280 NWYVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 339
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 433 ISKAKGQPREPVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTP 492
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 340 ISKAKGQPREPVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTP 399
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 493 PVLDSGDFLYSKLTVDKSRWQOGNVFSCSVWHEALHNNHYTKSLSPGK 544
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 400 PVLDSGDFLYSKLTVDKSRWQOGNVFSCSVWHEALHNNHYTKSLSPGK 451
   : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

```

RESULT 12  
US-08-887-352B-18  
; Sequence 18, Application US/08887352B  
; Patent No. 5994511  
; GENERAL INFORMATION:  
; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe  
; TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of  
; TITLE OF INVENTION: Improving Polypeptides  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/887,352B  
; FILING DATE: 03-Jul-1997  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Svoboda, Craig G.  
; REGISTRATION NUMBER: 39,044  
; REFERENCE/DOCKET NUMBER: P1123  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1489  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 451 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-887-352B-18

Query Match 42.3%; Score 1229.5; DB 2; Length 451;  
Best Local Similarity 61.7%; Pred. No. 2.3e-99;  
Matches 254; Conservative 16; Mismatches 53; Indels 89; Gaps 8;





```
QY 244 YVMKQGVDMKAMTSNVAQVCDNSDKFPV---YKYPGKGCPTLGDGDTLDLYDYPEEDT 300
Db 109 AVWQOGLTV-----TVSSASTKGPSVFLPAPSSKSTSGGTAALG---CLVKDYFPEPVT 159
QY 301 -----EGDKVRASTRVVKFPTKA-----
Db 160 VSMNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHNKPSNTKYDKKV 219
QY 319 -----RPHTCCPPCPAPEALGAPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKF 372
Db 220 EPKSCDKHTCCPPCPAPELGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKF 279
QY 373 NWYVDGVEVHNNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 432
Db 280 NWYVDGVEVHNNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 339
QY 433 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGYFSPDSIAVEMESNGQPENNYKTP 492
Db 340 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGYFSPDSIAVEMESNGQPENNYKTP 399
QY 493 PVLDSGDSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNYHTQKSLSPGK 544
Db 400 PVLDSGDSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNYHTQKSLSPGK 451

RESULT 15
US-09-109-207C-16
; Sequence 16, Application US/09109207C
; Patent No. 6172213
; GENERAL INFORMATION:
; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardiou, John Lowe
; TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of Improving Polypeptide
; FILE REFERENCE: P1123R1
; CURRENT APPLICATION NUMBER: US/09/109,207C
; CURRENT FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/051,554
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 44
; SEQ ID NO 16
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; NAME/KEY: Artificial
; LOCATION: 1-451
; OTHER INFORMATION: Heavy chain sequence derived from MAEL1
US-09-109-207C-16

Query Match 42.3%; Score 1229.5; DB 3; Length 451;
Best Local Similarity 61.7%; Pred. No. 2.3e-99;
Matches 254; Conservative 16; Mismatches 53; Indels 89; Gaps 8;

QY 188 ENLDPTLLQENSLYTIPIKGF-----GSHLLPFAFLHGNPMLCNCEILYFRWLQDIAENV 243
Db 74 DSKNTFYLQWNSLRAEDTAVYYCARGSH-----YFGHW-----HF 108
QY 244 YVMKQGVDMKAMTSNVAQVCDNSDKFPV---YKYPGKGCPTLGDGDTLDLYDYPEEDT 300
Db 109 AVWQOGLTV-----TVSSASTKGPSVFLPAPSSKSTSGGTAALG---CLVKDYFPEPVT 159
QY 301 -----EGDKVRASTRVVKFPTKA-----
Db 160 VSMNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHNKPSNTKYDKKV 219
QY 319 -----RPHTCCPPCPAPEALGAPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKF 372
Db 220 EPKSCDKHTCCPPCPAPELGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKF 279
QY 373 NWYVDGVEVHNNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 432
Db 280 NWYVDGVEVHNNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 339
```

```
QY 433 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGYFSPDSIAVEMESNGQPENNYKTP 492
Db 340 ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGYFSPDSIAVEMESNGQPENNYKTP 399
QY 493 PVLDSGDSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNYHTQKSLSPGK 544
Db 400 PVLDSGDSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNYHTQKSLSPGK 451

Search completed: October 14, 2003, 06:32:09
Job time : 26.266 secs
```

GenCore version 5.1.6  
Copy-right (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:16:09 ; Search time 24.6977 Seconds  
(without alignments)  
2067.627 Million cell updates/sec

Title: US-10-068-426-5

Perfect score: 2839

Sequence: 1 MPELLLLLLPLPLPHPHIC.....MHEALHHYTKSLSPGK 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1615.5	56.9	626	1 NBHUIA	platelet glycoprot
2	1205	42.4	330	1 GHUUA	Ig gamma-1 chain C
3	1203	42.4	374	2 S69339	Ig heavy chain V r
4	1197	42.2	255	4 S31866	Ig gamma-1 chain C
5	1152	40.6	234	2 PT0207	Ig gamma chain C r
6	1148	40.4	326	1 G2HU	Ig gamma-2 chain C
7	1145	40.3	377	2 A23511	Ig gamma-3 chain C
8	1143	40.3	377	2 A60764	Ig gamma-3 chain C
9	1134.5	40.0	327	1 C4HU	Ig gamma-4 chain C
10	1116.5	39.3	289	1 G3HUII	Ig gamma-4 heavy c
11	921	32.4	323	1 GHRB	Ig gamma chain C r
12	913	32.2	328	2 I47160	Ig gamma 2b chain
13	913	32.2	328	2 I47159	Ig gamma 2a chain
14	906	31.9	277	2 I47162	Ig gamma 4 chain c
15	893.5	31.5	328	2 I47158	Ig gamma 1 chain c
16	884.5	31.2	329	1 G2GP	Ig gamma-2 chain C
17	881	31.0	328	2 I47161	Ig gamma 3 chain c
18	857.5	30.2	470	2 S22080	Ig heavy chain pre
19	849.5	29.9	308	2 C30554	Ig heavy chain C r
20	849.5	29.9	472	2 S31459	Ig gamma-1 chain -
21	841	29.6	329	1 G3MSC	Ig gamma-3 chain C
22	838.5	29.5	444	2 PC4436	monoclonal antibod
23	834.5	29.4	326	2 PS0017	Ig gamma-1 chain C
24	830	29.2	398	1 G3MSM	Ig gamma-3 chain C
25	826.5	29.1	324	1 G1MS	Ig gamma-1 chain C
26	826.5	29.1	333	2 PS0018	Ig gamma-2b chain
27	821.5	28.9	329	2 S00847	Ig gamma-2c chain
28	821.5	28.9	393	1 G1MSM	Ig gamma-1 chain C
29	817.5	28.8	322	2 PS0019	Ig gamma-2a chain

## ALIGNMENTS

### RESULT 1

NBHUIA

platelet glycoprotein Ib alpha chain precursor - human

N:Alternate names: membrane glycoprotein Ib alpha chain

N:Contains: glycosialic

C:Species: Homo sapiens (man)

C>Date: 28-Dec-1987 #sequence revision 28-Dec-1987 #text change 22-Jun-1999

C/Accession: A94174; A60435; A94173; S16945; I55355; A27075; A27102

R/Lopez, J.A.; Chung, D.W.; Fujikawa, K.; Hagen, F.S.; Papayannopoulou, T.; Roth, G.J

Proc. Natl. Acad. Sci. U.S.A. 84, 5615-5619, 1987

A>Title: Cloning of the alpha-chain of human platelet glycoprotein Ib: a transmembran

A:Reference number: A94174; MUID:87289655; PMID:3303030

A:Accession: A94174

A:Molecule type: mRNA

A:Residues: 1-626 <LDP>

A:Cross-references: GB:J02940; NID:gi83499; PIDN:AAA52595.1; PID:G306793

R/Wicki, A.N.; Walz, A.; Gerber-Huber, S.N.; Wenger, R.H.; Vornhagen, R.; Clemetson, I

Thromb. Haemost. 61, 448-453, 1989

A>Title: Isolation and characterization of human blood platelet mRNA and construction

d cloning of a GPIb coding cDNA insert.

A:Reference number: A60435; MUID:90020160; PMID:2799758

A:Accession: A60435

A:Molecule type: mRNA

A:Residues: 207-467 <WIC>

R/Titani, K.; Takio, K.; Handa, M.; Ruggeri, Z.M.

Proc. Natl. Acad. Sci. U.S.A. 84, 5610-5614, 1987

A>Title: Amino acid sequence of the von Willebrand factor-binding domain of platelet

A:Reference number: A94173; MUID:87289654; PMID:3497398

A:Accession: A94173

A:Molecule type: protein

A:Residues: 17-315 <TIT>

R/Hess, D.; Schaller, J.; Rickli, E.E.; Clemetson, K.J.

Eur. J. Biochem. 199, 389-393, 1991

A>Title: Identification of the disulphide bonds in human platelet glycosialic.

A:Reference number: S16945; MUID:91301149; PMID:2070794

A:Accession: S16945

A>Status: preliminary

A:Molecule type: protein

A:Residues: 224-227;262-270;277-282 <HES>

R/Lopez, J.A.; Ludwig, E.H.; McCarthy, B.J.

J. Biol. Chem. 267, 10055-10061, 1992

A>Title: Polymorphism of human glycoprotein Ib alpha results from a variable number o

ations.

A:Reference number: I55355; MUID:92250564; PMID:1577776

A:Accession: I55355

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 412-427 <RES>

A:Cross-references: GB:S34436; NID:G249176; PIDN:AB22152.1; PID:G249177

A>Note: variant D

C:Comment: Glycoprotein Ib (GPIb), a surface membrane protein of platelets, participa

C:Comment: Platelet activation apparently involves disruption of the macromolecular c

C;Comment: Binding sites for von Willebrand factor and thrombin (the latter site with un  
C;Comment: Glycocalicin, which is approximately coextensive with the extracellular part  
C;Genetics:  
A;Gene: GDB:GPIBA; GPIB  
A;Map position: 17pter-17p12  
A;Cross-references: GDB:118806; OMIM:231200  
C;Complex: heterodimer with platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein  
C;Superfamily: platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein  
C;Keywords: blood coagulation; duplication; glycoprotein; platelet membrane; tandem rep  
F;1-16/Domain: signal sequence #status predicted <sig>  
F;17-626/Product: platelet glycoprotein Ib alpha chain #status predicted <MPT>  
F;48-71/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F;72-93/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F;94-116/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F;117-140/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
F;141-164/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
F;165-188/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
F;189-212/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
F;379-430/Region: proline/threonine-rich 9-residue repeats  
F;502-540/Domain: transmembrane #status predicted <TRM>  
F;541-626/Domain: intracellular #status predicted <INT>  
F;37,175/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F;308/Binding site: carbohydrate (Thr) (covalent) #status experimental

Query Match 56.9%; Score 1615.5; DB 1; Length 626;  
Best Local Similarity 84.2%; Pred. No. 1.2e-102;  
Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;

Qy	1	MPLLLLLLPSPLHPHPICEVSKVASHLEVNCNRLTALPPDLPKDTTLHLSENLLY	60
Db	1	MPLLLLLLPSPLHPHPICEVSKVASHLEVNCNRLTALPPDLPKDTTLHLSENLLY	60
Qy	61	TFSLATMPYTRLQNLDRCELTKLOVDGTLPLVGLTDLDSHNOQLSLGLGTLPLATV	120
Db	61	TFSLATMPYTRLQNLDRCELTKLOVDGTLPLVGLTDLDSHNOQLSLGLGTLPLATV	120
Qy	121	LDVSFNRLTSLPLGALRGLEQLYKLGKNEKLTLPGLLTPTPKLEKLSANNLTPL	180
Db	121	LDVSFNRLTSLPLGALRGLEQLYKLGKNEKLTLPGLLTPTPKLEKLSANNLTPL	180
Qy	181	AGLLNGLENLDTLLQENSLYTIKPGFFGSHLLPFAFLHGNPMVLCNEILYFRRWLQDNA	240
Db	181	AGLLNGLENLDTLLQENSLYTIKPGFFGSHLLPFAFLHGNPMVLCNEILYFRRWLQDNA	240
Qy	241	ENVVVKQGVDDVKAVTNSVAVQCDNSDKFVYKPKGCGPTLGDGDTLDLYYPEEDT	300
Db	241	ENVVVKQGVDDVKAVTNSVAVQCDNSDKFVYKPKGCGPTLGDGDTLDLYYPEEDT	300
Qy	301	EGDKVR-----PHTCP-----PCPA---PEALGAPSVFLFPKP--	330
Db	301	EGDKVRATRTVVKFPTKAHTTPMGFLYSWSASTASLDQMPSSLHTPTQSTKEQTTPPRWT	360
Qy	331	PKDTL-----MISTPEVT	344
Db	361	PNFTLHMESTFSPKST	379

RESULT 2  
GHU

Ig gamma-1 chain C region - human  
C;Species: Homo sapiens (man)  
C;Date: 31-Jan-1981 #sequence revision 18-Aug-1982 #text change 16-Jul-1999  
C;Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146  
R;Ellison, J.W.; Berson, B.J.; Hood, L.E.  
Nucleic Acids Res. 10, 4071-4079, 1982  
A;Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.  
A;Reference number: A93433; MUID:82274238; PMID:6287432  
A;Accession: A93433  
A;Molecule type: DNA  
A;Residues: 1-330 <ELL>  
A;Cross-references: EMBL:Z17370  
A;Note: This sequence has the Gim(17) allotypic marker, 97-Lys, and the Gim(1) markers,  
A;Note: Lys-330 is removed after translation

R;Harris, L.J.  
Submitted to the EMBL Data Library, October 1992  
A;Reference number: S33904  
A;Accession: S36861  
A;Molecule type: DNA  
A;Residues: 2-330 <HAR>  
A;Cross-references: EMBL:Z17370  
R;Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.  
Cell 29, 671-679, 1982  
A;Title: Structure of human immunoglobulin gamma genes: implications for evolution of  
A;Reference number: S33887; MUID:83001943; PMID:6811139  
A;Accession: S33887  
A;Molecule type: DNA  
A;Residues: 88-113/235-330 <TAK>  
A;Cross-references: EMBL:Z17370  
R;Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman  
Biochemistry 9, 3161-3170, 1970  
A;Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid seq  
A;Reference number: A90563; MUID:71064024; PMID:5489771  
A;Contents: myeloma protein Eu  
A;Accession: B90563  
A;Molecule type: protein  
A;Residues: 1-96,'R',98-135 <CUN>  
R;Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.  
Biochemistry 9, 3171-3181, 1970  
A;Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid se  
A;Reference number: A90564; MUID:71064025; PMID:5530842  
A;Contents: Eu  
A;Accession: A90564  
A;Molecule type: protein  
A;Residues: 136-154,'O',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',2  
A;Note: this sequence has the Gim(non-1) markers, 239-Glu and 241-Met  
R;Ponstingl, H.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976  
A;Title: Die Primärstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Ni  
igen Primärstruktur.  
A;Reference number: A91668; MUID:77070269; PMID:826475  
A;Contents: myeloma protein Nie  
A;Accession: B91668  
A;Molecule type: protein  
A;Residues: 1-34,'Q',36-96,'K',98-115,'Q',117-197,'D',199-238,'D',240,'L',242-268,'E'  
A;Note: this sequence has the Gim(17) and Gim(1) markers  
R;Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983  
A;Title: Die Primärstruktur des kristallisierbaren monoklonalen Immunglobulins IgG1  
A;Reference number: A91723; MUID:83289131; PMID:6884994  
A;Contents: myeloma protein KOL; disulfide bonds  
A;Accession: A91723  
A;Molecule type: protein  
A;Residues: 1-96,'R',98-197,'D',199-238,'E',240,'M',242-266,'D',268-271,'D',273-330 <  
A;Note: this sequence has the Gim(3) and Gim(non-1) markers  
R;Gall, W.E.; Edelman, G.M.  
Biochemistry 9, 3188-3196, 1970  
A;Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disul  
A;Reference number: A90565; MUID:71064027; PMID:4923144  
A;Contents: annotation; disulfide bonds  
R;Breker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976  
A;Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunog  
enonide cleavage products, and the disulfide bridges.  
A;Reference number: A91667; MUID:77070267; PMID:1002129  
A;Contents: annotation; disulfide bonds  
C;Genetics:  
A;Gene: GDB:IGHG1  
A;Cross-references: GDB:120085; OMIM:147100  
A;Map position: 14q32.33-14q32.33  
A;Introns: 9/1; 114/1; 224/1  
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (h  
ain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F;20-85/Domain: immunoglobulin homology <IM1>

```

QY 426 QPREQVVTLLPSREEMTKQVSLTCLVKGFYPSDIAVEMESNGOPENNYKTTTPVLDS 485
DB 269 QPREQVVTLLPSREEMTKQVSLTCLVKGFYPSDIAVEMESNGOPENNYKTTTPVLDS 328
QY 486 GSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTQKLSLSPOK 531
DB 329 GSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTQKLSLSPOK 374

RESULT 4
S31866
Ig gamma-1 chain C region - synthetic
C:Species: synthetic
A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli
C:Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000
C:Accession: S31866
R:Filpula, D.
submitted to the EMBL Data Library, February 1993
A:Description: Screening method for protein-protein interactions of cloned gene products
A:Reference number: S31866
A:Accession: S31866
A:Molecule type: mRNA
A:Residues: 1-255 <FIL>
A:Cross-references: EMBL:X70421, NID:g33068, PIDN:CAA49866.1, PID:g33069
C:Keywords: immunoglobulin
F:1-22/Region: Escherichia coli outer membrane protein A precursor
F:23-255/Region: human Ig gamma-1 chain C region

Query Match 42.2%; Score 1197; DB 4; Length 255;
Best Local Similarity 96.9%; Pred.No. 1.4e-74;
Matches 219; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 306 RHTCPPCPAPEALGAPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 365
DB 30 KHTCTCPCPAPELLGSPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 89

QY 366 VEVHNAKTPREEQNSTYRVVSVLTVTLHQDLNGKEYCKVSKNPKALPVIETISKAG 425
DB 90 VEVHNAKTPREEQNSTYRVVSVLTVTLHQDLNGKEYCKVSKNPKALPVIETISKAG 149

QY 426 QPREQVVTLLPSREEMTKQVSLTCLVKGFYPSDIAVEMESNGOPENNYKTTTPVLDS 485
DB 150 QPREQVVTLLPSREEMTKQVSLTCLVKGFYPSDIAVEMESNGOPENNYKTTTPVLDS 209

QY 486 GSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTQKLSLSPOK 531
DB 210 GSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTQKLSLSPOK 255

RESULT 5
PT0207
Ig gamma chain C region - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C:Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 16-Jul-1999
C:Accession: PT0207
R:Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.
Mol. Immunol. 28, 319-322, 1991
A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.
A:Reference number: PT0207; MUID:91287716; PMID:2062315
A:Accession: PT0207
A:Molecule type: mRNA
A:Residues: 1-234 <EHR>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:48-117/Domain: immunoglobulin homology <IMM>

Query Match 40.6%; Score 1152; DB 2; Length 234;
Best Local Similarity 93.8%; Pred.No. 1.5e-71;
Matches 213; Conservative 3; Mismatches 5; Indels 6; Gaps 1;

QY 304 KVRP-----HTCPPCPAPEALGAPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEV 357
DB 8 KVRKSCDTHTCPCAPPELIGGSPVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEV 57

```

QY 358 KFNYYVGVGVHNAKTPREEOYNSTYRVVSVLTVLHODWLNCKEYKCKVSNKALPVPIE 417  
 DB 68 KFNYYVGVGVHNAKTPREEQNSTYRVVSVLTVLHODWLNCKEYKCKVSNKALPVPIE 127  
 QY 418 KTSKAKGQPREQVYTLTPSRDELTQVSLTCLVKGFYPSDIAVEMESNGOPENNYKT 477  
 DB 128 KTSKAKGQPREQVYTLTPSRDELTQVSLTCLVKGFYPSDIAVEMESNGOPENNYKT 187  
 QY 478 TTPVLSDSGFFLYSKLTVDKSRWQGNVPSCSMHEALHNHYTQKS 524  
 DB 188 TTPVLSDSGFFLYSKLTVDKSRWQGNVPSCSMHEALHNHYTQKS 234

RESULT 6  
 G2HU  
 Ig gamma-2 chain C region - human  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Apr-1981 #sequence revision 13-Jun-1983 #text\_change 21-Jul-2000  
 C:Accession: A93906; A92809; A90752; A93132; A02148  
 R:Ellison, J.; Hood, L.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982  
 A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con  
 A:Reference number: A93906; MUID:82197621; PMID:6804948  
 A:Accession: A93906  
 A:Molecule type: DNA  
 A:Residues: 1-326 <ELL>  
 A:Cross-references: GB:V00554; GB:J00230; NID:g32759; PIDN:CABS8438.1; PID:g6066056  
 A:Note: Lys-326 is probably removed posttranslationally  
 R:Wang, A.C.; Tung, E.; Fudenberg, H.H.  
 J. Immunol. 125, 1048-1054, 1980  
 A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and e  
 A:Reference number: A92809; MUID:81007873; PMID:6774012  
 A:Contents: myeloma protein T11  
 A:Accession: A92809  
 A:Molecule type: protein  
 A:Residues: 1-19, 'Q', 21-57, 'Z', 59, 'A', 61-193, 'D', 195-325 <WAN>  
 A:Note: Trp-156 is at or near the complement-binding site  
 R:Connell, G.E.; Parr, D.M.; Hofmann, T.  
 Can. J. Biochem. 57, 758-767, 1979  
 A:Title: The amino acid sequences of the three heavy chain constant region domains of a  
 A:Reference number: A90752; MUID:80001357; PMID:113060  
 A:Contents: myeloma protein Zie  
 A:Accession: A90752  
 A:Molecule type: protein  
 A:Residues: 1-24, 'E', 26-57, 'EV', 60-85:132-171, 'ZZZ', 175, 'B', 177-193, 'D', 195-196, 'Q', 198-  
 A:Note: this sequence has since been revised  
 R:Hofmann, T.; Parr, D.M.  
 Mol. Immunol. 16, 923-925, 1979  
 A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g  
 A:Reference number: A93132; MUID:80114419; PMID:118920  
 A:Contents: Zie  
 A:Accession: A93132  
 A:Molecule type: protein  
 A:Residues: 238-275 <HOF>  
 R:Hofmann, T.; Parr, D.M.  
 submitted to the Atlas, March 1980  
 A:Reference number: A94591  
 A:Contents: annotation; Zie, revisions to residues 25, 59, 60, and 264-268  
 A:Note: the revised sequence differs from that shown in having 60-Ala and in the amidati  
 ned  
 R:Milstein, C.; Frangione, B.  
 Biochem. J. 121, 217-225, 1971  
 A:Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.  
 A:Reference number: A90253; MUID:72033500; PMID:4940472  
 A:Contents: annotation; myeloma protein Sa, disulfide bonds  
 R:Frangione, B.; Milstein, C.; Pink, J.R.L.  
 Nature 221, 145-148, 1969  
 A:Title: Structural studies of immunoglobulin G.  
 A:Reference number: A93157; MUID:69064124; PMID:5782707  
 A:Contents: annotation; Sa, disulfide bonds  
 C:Genetics:  
 A:Gene: GDB:IGHG2

A:Cross-references: GDB:119338; OMIM:147110  
 A:Map position: 14q32.33-14q32.33  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light ( chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IM1>  
 F:133-202/Domain: immunoglobulin homology <IM2>  
 F:239-306/Domain: immunoglobulin homology <IM3>  
 F:14/Disulfide bonds: interchain (co light chain) #status experimental  
 F:27-83,140-200,246-304/Disulfide bonds: #status experimental  
 F:102,103,106,109/Disulfide bonds: interchain (co heavy chain) #status experimental  
 F:176/Binding site: carbohydrate (Asn) #status predicted

Query Match 40.4%; Score 1148; DB 1; Length 326;  
 Best Local Similarity 72.2%; Pred. No. 4.4e-71;  
 Matches 226; Conservative 18; Mismatches 33; Indels 36; Gaps 6;

QY 253 KAVTSNVAHVQCDNSDKFP-----VYKYP-----GKGCPYL----- 283  
 DB 16 RSTSESTAALGCLVKDYFPEPTVYSWNSGALTSGVHTFPAVLQSSGLYSLSSVVPSSN 75  
 QY 284 -GDBGDTLDYDYEEDTEGDKVRPH-----TCPPCPAPEALGAPSVLFPKPKDTLMIS 338  
 DB 76 FGTQYTCNVDPKHP-SNTKVDKTVVERKCCVPCPPAVAG-PSVFLFPKPKDTLMIS 133  
 QY 339 RTPVTCVVVDVSHEDPEVKFNWYVDGVEFNKATKPREEOYNSTYRVVSVLTVLHODWL 398  
 DB 134 RTPVTCVVVDVSHEDPEVKFNWYVDGVEFNKATKPREEOYNSTYRVVSVLTVLHODWL 193  
 QY 399 NGKEYKCKVSNKALPVPKEITISKAKGQPREPQVYTLTPPSREEMTKNQVSLTCLVKGFYP 458  
 DB 194 NGKEYKCKVSNKALPVPKEITISKAKGQPREPQVYTLTPPSREEMTKNQVSLTCLVKGFYP 253  
 QY 459 SDIAVEMESNGOPENNYKTTPPVLSDSGFFLYSKLTVDKSRWQGNVPSCSMHEALHN 518  
 DB 254 SDIAVEMESNGOPENNYKTTPPVLSDSGFFLYSKLTVDKSRWQGNVPSCSMHEALHN 313  
 QY 519 HYTKQSLSLSPCK 531  
 DB 314 HYTKQSLSLSPCK 326

RESULT 7  
 A23511  
 Ig gamma-3 chain C region (allotype G3m(b)) - human  
 C:Species: Homo sapiens (man)  
 C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 23-Jul-1999  
 C:Accession: A23511  
 R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.  
 Nucleic Acids Res. 14, 1779-1789, 1986  
 A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene:  
 A:Reference number: A23511; MUID:86148507; PMID:3081877  
 A:Accession: A23511  
 A:Molecule type: DNA  
 A:Residues: 1-377 <HUC>  
 A:Cross-references: GB:X03604; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID:g577056  
 C:Genetics:  
 A:Gene: GDB:IGHG3  
 A:Cross-references: GDB:119339; OMIM:147120  
 A:Map position: 14q32.33-14q32.33  
 A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 40.3%; Score 1145; DB 2; Length 377;  
 Best Local Similarity 64.4%; Pred. No. 8.7e-71;  
 Matches 230; Conservative 17; Mismatches 56; Indels 54; Gaps 7;

QY 194 LIQENSLY-----TIPKGFQSHLLPFAFLHGNPWLNCCELLYPRRWLQDAENYVWK 247  
 DB 56 VLQSSGLYSLSSVWTVPSSSLGT-----QTYTCN-----VNHKPSNTKYDK 96









C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match	31.5%	Score	893.5	DB 2	Length	328			
Best Local Similarity	68.7%	Pred. No. 1.le-53							
Matches	167	Conservative	29	Mismatches	40	Indels	7	Gaps	3

  

QY	295	YPEDTEGDK----	VRPHTCPGPCPAPENALCAPSVFLFPPKPKDTLMISRTPEVTCVWVDV	350
DB	87	HPATTTKVDKRVGIHQFQTCPCPGCEVAG-PSVFIFPPKPKDTLMISQTPETVCVWVDV	145	
QY	351	SHEDPEVKEVWYVDGVEVHNAKTPREEQVNSTYRVSVLTVLHQDWLNGKEYKCKVSNK	410	
DB	146	SKEHAEVQFSWYVDGVEVHTAETRPKEEQFNSTYRVSVLPVQHODWLKGEKCKVNV	205	
QY	411	ALPVPTEKTISKAGQPREPOVYTLPPSREEMTKNQVSLTCLVKGYFSPSDIAVEWESNGQ	470	
DB	206	DLPAPIRTTISKAGQPREQVYTLPPPAEELSRSKVTLTCLVIGFYPPDIHVEWKSNGQ	265	
QY	471	--PENNYKTTTPVLDSDGSFELYSKLTVDKSRWQQGVFSCSYMHEALHNHYTQKSLS	528	
DB	266	PEPENTYRTTTPPOODVDGTFFLYSKLAVDKARWDHGDGKFECAVMHEALHNHYTQKS	325	
QY	529	PGK	531	
DB	326	QK	328	

Search completed: October 14, 2003, 06:17:18  
Job time : 25.6977 secs

Result No.	Query			ID	Description
	Score	Match	Length		
1	1615.5	56.9	626	1	GPBA_HUMAN
2	1205	42.4	330	1	GCL_HUMAN
3	1148	40.4	326	1	GC2_HUMAN
4	1134.5	40.0	327	1	GC4_HUMAN
5	1121.5	39.5	290	1	GC3_HUMAN
6	921	32.4	323	1	GC_RABIT
7	884.5	31.2	329	1	GC2_CAVPO
8	841	29.6	329	1	GC3_MOUSE
9	834.5	29.4	326	1	GCL_RAT
10	830	29.2	398	1	GC3M_MOUSE
11	826.5	29.1	324	1	GCL_MOUSE
12	826.5	29.1	333	1	GCB_RAT
13	821.5	28.9	329	1	GCC_RAT
14	821.5	28.9	393	1	GC1M_MOUSE
15	817.5	28.8	332	1	GCA_RAT
16	809	28.5	330	1	GCAA_MOUSE
17	804.5	28.3	335	1	GCAH_MOUSE
18	804	28.3	399	1	GCAM_MOUSE
19	770	27.1	336	1	GCB_MOUSE
20	765	26.9	405	1	GCBM_MOUSE
21	363	12.8	421	1	EPC_MOUSE
22	360	12.7	391	1	MUCB_HUMAN
23	359.5	12.7	429	1	EPC_RAT
24	359	12.6	454	1	MUC_HUMAN
25	352	12.4	455	1	MUC_MOUSE
26	342	12.0	476	1	MUCM_MOUSE
27	341	12.0	428	1	EPC_HUMAN
28	338	11.9	458	1	MUC_RABIT
29	331.5	11.7	450	1	MUC_CANFA
30	329	11.6	454	1	MUC_MESAU
31	328	11.6	479	1	MUCM_RABIT
32	322	11.3	457	1	MUC_SUNMU
33	301	10.6	447	1	A2GL_HUMAN

RT "Identification of the disulphide bonds in human platelet  
 RL glycofibrinogen";  
 RN Sur. J. Biochem. 199:389-393(1991).  
 RP [7]  
 RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).  
 RX MEDLINE=22171415; PubMed=12183630;  
 RA Huizinga E.G., Tsuji S., Romijn R.A., Schiphorst M.E., de Groot P.G.,  
 RA Sixma J.J., Gros P.;  
 RT "Structures of glycoprotein Ib alpha and its complex with von  
 RT Willebrand factor A1 domain";  
 RL Science 297:1176-1179(2002).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 17-304, AND SULFATION OF  
 RP TYR-292; TYR-294 AND TYR-295.  
 RX MEDLINE=22217968; PubMed=12087105;  
 RA Uff S., Clemenson J.M., Harrison T., Clemenson K.J., Emsley J.;  
 RT "Crystal structure of the platelet glycoprotein Ib alpha N-terminal  
 RT domain reveals an unmasking mechanism for receptor activation";  
 RL J. Biol. Chem. 277:35657-35663(2002).  
 RN [9]  
 RP VARIANT SJBA MET-161.  
 RX MEDLINE=92265982; PubMed=1586750;  
 RA Murata M., Furihata K., Ishida F., Russell S.R., Ware J.,  
 RA Ruggeri Z.M.;  
 RT "Genetic and structural characterization of an amino acid dimorphism  
 RT in glycoprotein Ib alpha involved in platelet transfusion  
 RT refractoriness";  
 RL Blood 79:3086-3090(1992).  
 RN [10]  
 RP VARIANT BSS PHE-73.  
 RX MEDLINE=92110577; PubMed=1730088;  
 RA Miller J.L., Lyle V.A., Cunningham D.;  
 RT "Mutation of leucine-57 to phenylalanine in a platelet glycoprotein:  
 RT Ib alpha leucine tandem repeat occurring in patients with an  
 RT autosomal dominant variant of Bernard-Soulier disease";  
 RL Blood 79:439-446(1992).  
 RN [11]  
 RP VARIANT BSS VAL-172.  
 RX MEDLINE=93388951; PubMed=7690774;  
 RA Ware J., Russell S.R., Marchese P., Murata M., Mazzucato M.,  
 RA De Marco L., Ruggeri Z.M.;  
 RT "Point mutation in a leucine-rich repeat of platelet glycoprotein Ib  
 RT alpha resulting in the Bernard-Soulier syndrome";  
 RL J. Clin. Invest. 92:1213-1220(1993).  
 RN [12]  
 RP VARIANT BSS SER-225.  
 RX MEDLINE=95118882; PubMed=7819107;  
 RA Simsek S., Norris P., Lozano M., Pico M., von Dem Borne A.E.G.K.,  
 RA Ribera A., Gallardo D.;  
 RT "Cys209 Ser mutation in the platelet membrane glycoprotein Ib alpha  
 RT gene is associated with Bernard-Soulier syndrome";  
 RL Br. J. Haematol. 88:839-844(1994).  
 RN [13]  
 RP VARIANT PSEUDO-VWD VAL-249.  
 RX MEDLINE=91271273; PubMed=2052556;  
 RA Miller J.L., Cunningham D., Lyle V.A., Finch C.N.;  
 RT "Mutation in the gene encoding the alpha chain of platelet  
 RT glycoprotein Ib in platelet-type von Willebrand disease";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:4761-4765(1991).  
 RN [14]  
 RP VARIANT PSEUDO-VWD VAL-249.  
 RX MEDLINE=93253059; PubMed=8486780;  
 RA Murata M., Russell S.R., Ruggeri Z.M., Ware J.;  
 RT "Expression of the phenotypic abnormality of platelet-type von  
 RT Willebrand disease in a recombinant glycoprotein Ib alpha fragment";  
 RL J. Clin. Invest. 91:2133-2137(1993).  
 RN [15]  
 RP VARIANT PSEUDO-VWD VAL-255.  
 RX MEDLINE=93214031; PubMed=8384898;  
 RA Russell S.D., Roth G.J.;  
 RT "Pseudo-von Willebrand disease: a mutation in the platelet  
 RT glycoprotein Ib alpha gene associated with a hyperactive surface  
 RT receptor";

Blood 81:1787-1791(1993).  
 [16]  
 VARIANT BSS LEU-195 DEL.  
 MEDLINE=95178321; PubMed=7873390;  
 de la Salle C., Baas M.-J., Lanza F., Schwartz A., Hanau D.,  
 Chevalier J., Gachet C., Briquel M.-E., Cazenave J.-P.;  
 "A three-base deletion removing a leucine residue in a leucine-rich  
 repeat of platelet glycoprotein Ib alpha associated with a variant of  
 Bernard-Soulier syndrome (Nancy 1)";  
 Br. J. Haematol. 89:386-396(1995).  
 [17]  
 VARIANT BSS ARG-81.  
 MEDLINE=98303759; PubMed=9639514;  
 Kenny D., Jonsson O.G., Moratsek P.A., Montgomery R.R.;  
 "Naturally occurring mutations in glycoprotein Ib alpha that result in  
 defective ligand binding and synthesis of a truncated protein";  
 Blood 92:175-183(1998).  
 [18]  
 VARIANTS HIS-72 AND MET-161.  
 MEDLINE=99318093; PubMed=10391209;  
 Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 Lander E.S.;  
 "Characterization of single-nucleotide polymorphisms in coding regions  
 of human genes";  
 Nat. Genet. 22:231-238(1999).  
 [19]  
 ERRATUM.  
 Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,  
 Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 Lander E.S.;  
 Nat. Genet. 23:373-373(1999).  
 [20]  
 VARIANT BSS PRO-145.  
 MEDLINE=99189763; PubMed=10089893;  
 Koskela S., Partanen J., Salmi T.T., Kekomaki R.;  
 "Molecular characterization of two mutations in platelet glycoprotein  
 (GP) Ib alpha in two Finnish Bernard-Soulier syndrome families";  
 Eur. J. Haematol. 62:160-168(1999).  
 CC -!- FUNCTION: GP-Ib, a surface membrane protein of platelets,  
 CC participates in the formation of platelet plugs by binding to the  
 CC A1 domain of von Willebrand factor, which is already bound to the  
 CC subendothelium.  
 CC -!- SUBUNIT: Heterodimer composed of GP-Ib alpha and beta; disulfide  
 CC linked. GP-IX is complexed with the GP-Ib heterodimer via a non  
 CC covalent linkage.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- PTM: Glycosylation, which is approximately coextensive with the  
 CC extracellular part of the molecule, is cleaved off by calpain  
 CC during platelet lysis.  
 CC -!- POLYMORPHISM: Position 161 is associated with platelet-specific  
 CC alloantigen Siba. Siba(-) has Thr-161 and Siba(+) has Met-161.  
 CC Siba is involved in neonatal alloimmune thrombocytopenia (NATP).  
 CC -!- DISEASE: Defects in GPIBA are one of the causes of Bernard-Soulier  
 CC syndrome (BSS). BSS patients have unusually large platelets and  
 CC have a clinical bleeding tendency.  
 CC -!- DISEASE: Defects in GPIBA are one of the causes of von Willebrand  
 CC disease (vWD) known as platelet-type von Willebrand disease or  
 CC pseudo-von Willebrand disease. This autosomal dominant bleeding  
 CC disorder is caused by an increased affinity of GP-Ib for soluble  
 CC vWF resulting in impaired hemostatic function due to the removal  
 CC of vWF from the circulation.  
 CC -!- MISCELLANEOUS: Platelet activation apparently involves disruption  
 CC of the macromolecular complex of GP-Ib with the platelet  
 CC glycoprotein IX (GP-IX) and dissociation of GP-Ib from the actin-  
 CC binding protein.  
 CC -!- MISCELLANEOUS: Binding sites for von Willebrand factor and  
 CC thrombin (the latter site with unknown function) are in the amino-  
 CC terminal part of the molecule.  
 CC -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.

Query Match 56.9%; Score 1615.5; DB 1; Length 626;  
Best Local Similarity 84.2%; Pred. No. 8.8e-108;  
Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;

QY 1 MPELLLLLLLPSPLPHPICEVSKVASHLEVNCCKRNLTPALPDLPKOTITLHLSNLLY 60  
DB 1 MPELLLLLLLPSPLPHPICEVSKVASHLEVNCCKRNLTPALPDLPKOTITLHLSNLLY 60

QY 61 TFSLATLMPYTRLTQNLORCELTKLQVDTLPVLGTLDSLHNSQLSPLGLQOTIPALTV 120  
DB 61 TFSLATLMPYTRLTQNLORCELTKLQVDTLPVLGTLDSLHNSQLSPLGLQOTIPALTV 120

QY 121 LDVSFNRLTSLPGALRGELGELQELYLKGNELKTLPPGLLTPPKLEKLSLNNLITELP 180  
DB 121 LDVSFNRLTSLPGALRGELGELQELYLKGNELKTLPPGLLTPPKLEKLSLNNLITELP 180

QY 181 AGLLNGLENDTLQLQENSLYTIKPGFFGSHLLPFAFLHGNPWLNCNCEILYFRRMLQDNA 240  
DB 181 AGLLNGLENDTLQLQENSLYTIKPGFFGSHLLPFAFLHGNPWLNCNCEILYFRRMLQDNA 240

QY 241 ENYVYWKQVYKAVTSNVASVOCNSDKFPYKYPKSCPTLGDGDTLDLYDYYPEEDT 300  
DB 241 ENYVYWKQVYKAVTSNVASVOCNSDKFPYKYPKSCPTLGDGDTLDLYDYYPEEDT 300

QY 301 EGDQVR-----PHTCP-----PCPA---PEALGAPSVFLPPPK-- 330  
DB 301 EGDQVRATRVVKEFKKANTPGLFYSWSTASLDSQMPSSLUHPTQESTKEQTTPPRWT 360

QY 331 PKDTL-----MISRTPEVT 344  
DB 361 PNFTLHMSITFSKTPKST 379

## RESULT 2

GC1\_HUMAN STANDARD; PRT; 330 AA.  
ID GC1\_HUMAN  
AC P01857;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Ig gamma-1 chain C region.  
GN IGHG1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=82274238; PubMed=6287432;  
RA Ellison J.W., Berson B.J., Hood L.E.;  
RT "The nucleotide sequence of a human immunoglobulin C gamma1 gene.";  
RL Nucleic Acids Res. 10:4071-4079(1982).  
RN [2]  
RP SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).  
RX MEDLINE=71064024; PubMed=5489771;  
RA Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,  
RA Wexdal M.J., Edelman G.M.;  
RT "The covalent structure of a human gamma G-immunoglobulin. VII. Amino  
RL acid sequence of heavy-chain cyanogen bromide fragments H1-H4.";  
RL Biochemistry 9:3161-3170(1970).  
RN [3]  
RP SEQUENCE OF 136-329 (EU).  
RX MEDLINE=71064025; PubMed=5530842;  
RA Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,  
RA Edelman G.M.;  
RT "The covalent structure of a human gamma G-immunoglobulin. 8. Amino  
RL acid sequence of heavy-chain cyanogen bromide fragments H5-H7.";  
RL Biochemistry 9:3171-3181(1970).  
RN [4]  
RP SEQUENCE (MYELOMA PROTEIN NIE).  
RX MEDLINE=77070269; PubMed=826475;  
RA Ponstingl H., Hilschmann N.;

RT "The rule of antibody structure. The primary structure of a  
RT monoclonal IgG1 immunoglobulin (myeloma protein NIE). III. The  
RT chymotryptic peptides of the H-chain, alignment of the tryptic  
RT peptides and discussion of the complete structure.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).  
RN [5]  
RP SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.  
RX MEDLINE=83289131; PubMed=6884994;  
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;  
RT "Three-dimensional structure determination of antibodies. Primary  
RT structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).  
RN [6]  
RP DISULFIDE BONDS.  
RX MEDLINE=71064027; PubMed=4923144;  
RA Gall W.E., Edelman G.M.;  
RT "The covalent structure of a human gamma G-immunoglobulin. X.  
RT Intrachain disulfide bonds.";  
RL Biochemistry 9:3188-3196(1970).  
RN [7]  
RP DISULFIDE BONDS.  
RX MEDLINE=77070267; PubMed=1002129;  
RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;  
RT "Rule of antibody structure. The primary structure of a monoclonal  
RT IgG1 immunoglobulin (myeloma protein NIE). I: Purification and  
RT characterization of the protein, the L- and H-chains, the  
RT cyanogen bromide cleavage products, and the disulfide bridges.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
RX MEDLINE=81208100; PubMed=7236608;  
RA Deisenhofer J.;  
RT "Crystallographic refinement and atomic models of a human Fc fragment  
RT aureus at 2.9 and 2.8-A resolution.";  
RL Biochemistry 20:2361-2370(1981).  
CC -!- MISCELLANEOUS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE  
CC GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)  
CC MARKER & THE GIM (NON-1) MARKERS.  
CC -!- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF  
CC 35,116,198,269 & 272.  
CC -!- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES  
CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES  
CC 268-272.  
CC -!- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF  
CC RESIDUES 198, 267&272.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC  
CC EMBL; J00228; AAC82527.1; ALT\_INIT.  
DR PIR; A93433; GHU.  
DR PDB; 1FC1; 15-JUL-92.  
DR PDB; 1FC2; 15-JUL-92.  
DR PDB; 1AJ7; 12-NOV-97.  
DR PDB; 1DSB; 09-FEB-00.  
DR PDB; 1DS1; 09-FEB-00.  
DR PDB; 1D6V; 04-OCT-00.  
DR PDB; 1DN2; 17-MAY-00.  
DR PDB; 1S4K; 06-JUN-01.  
DR PDB; 1FCC; 20-JUL-95.  
DR PDB; 1H2H; 12-JUN-02.  
DR PDB; 1I7Z; 08-AUG-01.  
DR PDB; 1I1S; 16-MAY-01.  
DR PDB; 1I1X; 16-MAY-01.  
DR PDB; 1I6X; 10-APR-02.  
DR PDB; 2RCS; 12-NOV-97.  
DR GenBank; HGNC:5525; IGHG1.

DR MIM; 147100; --  
 DR GO; GO:0005624; C-membrane fraction; NAS.  
 DR GO; GO:0003823; F-antigen binding activity; TAS.  
 DR GO; GO:0006955; P-immune response; NAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig\_c1.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00407; Igcl; 2.  
 DR PROSITE; PS00835; IG\_LIKE; 3.  
 DR PROSITE; PS00290; IG\_MHC; 2.  
 DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 3D-structure.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 98 CH1.  
 FT DOMAIN 99 110 HINGE.  
 FT DOMAIN 111 223 CH2.  
 FT DOMAIN 224 330 CH3.  
 FT DISULFID 27 83  
 FT DISULFID 103 103 INTERCHAIN (WITH LIGHT CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN).  
 FT DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).  
 FT DISULFID 144 204  
 FT DISULFID 250 308  
 FT CARBOHYD 180 180  
 FT MOD\_RES 330 330  
 FT VARIANT 97 97  
 FT VARIANT 239 239  
 FT VARIANT 241 241  
 FT STRAND 122 126  
 FT HELIX 130 134  
 FT TURN 136 137  
 FT STRAND 141 147  
 FT TURN 157 162  
 FT STRAND 163 164  
 FT STRAND 165 166  
 FT TURN 168 171  
 FT STRAND 176 179  
 FT TURN 180 181  
 FT STRAND 182 190  
 FT HELIX 193 197  
 FT TURN 198 199  
 FT STRAND 202 207  
 FT TURN 209 210  
 FT STRAND 215 219  
 FT STRAND 227 227  
 FT STRAND 230 234  
 FT HELIX 238 242  
 FT STRAND 245 256  
 FT STRAND 260 265  
 FT STRAND 270 270  
 FT STRAND 274 276  
 FT STRAND 280 281  
 FT TURN 283 284  
 FT STRAND 287 296  
 FT HELIX 297 301  
 FT TURN 302 303  
 FT STRAND 305 312  
 FT TURN 313 314  
 FT TURN 316 317  
 FT STRAND 320 325  
 SQ SEQUENCE 330 AA; 36106 MW; 3770BE106C2FA33D CRC64;  
 Query Match 42.4%; Score 1205; DB 1; Length 330;  
 Best Local Similarity 73.7%; Pred. No. 7.5e-79;  
 Matches 235; Conservative 10; Mismatches 30; Indels 44; Gaps 5;  
 QY 253 KAVTSNVAQCDNSDKFP-----VYKYPGKGCPTLGDGEG----- 287  
 DB 16 KSTSGGTAALGCLVKDYFPEPTVTSWNSGALTSGVHTFPA-----VLSGGIYSLSSVVT 71

QY 288 -----DTDLY-----DYYPEDETEGDKVRP-----HTCPPCPAPEALCAPSVFLFPPKPK 332  
 DB 72 PSSSLGTQTYICNVNHPKSNTKVDKVEPKSCDKTHCTCPCPAPPELLGGPSVFLFPPKPK 131  
 QY 333 DTLMSIRTPETVCVVVSHEDPEVKFNWYVDGVEVFNNAKTKPREEQNSYTRVVSULTV 392  
 DB 132 DTLMSIRTPETVCVVVSHEDPEVKFNWYVDGVEVFNNAKTKPREEQNSYTRVVSULTV 191  
 QY 393 LHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRREEMTKNQVSLTCL 452  
 DB 192 LHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCL 251  
 QY 453 VKGYFPDSIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSSVM 512  
 DB 252 VKGYFPDSIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSSVM 311  
 QY 513 HEALNHYTQKSLSPGK 531  
 DB 312 HEALNHYTQKSLSPGK 330  
 RESULT 3  
 GC2\_HUMAN  
 ID GC2\_HUMAN STANDARD; PRT; 326 AA.  
 AC P01859;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ig gamma-2 chain C region.  
 GN IGHG2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE OF 2-326 FROM N.A.  
 RX MEDLINE=82197621; PubMed=6804948;  
 RA Ellison J.W., Hood L.E.;  
 RT "Linkage and sequence homology of two human immunoglobulin gamma  
 RT heavy chain constant region genes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).  
 RN [2]  
 RP SEQUENCE OF 88-115 FROM N.A.  
 RC TISSUE=Fetal liver;  
 RX MEDLINE=83001943; PubMed=6811139;  
 RA Takahashi N., Ueda S., Obata M., Nikaide T., Nakai S., Honjo T.;  
 RT "Structure of human immunoglobulin gamma genes: implications for  
 RT evolution of a gene family.";  
 RL Cell 29:671-679(1982).  
 RN [3]  
 RP SEQUENCE OF 99-177 AND 310-326 FROM N.A.  
 RC TISSUE=Fetal liver;  
 RX MEDLINE=84235992; PubMed=6329676;  
 RA Krawinkel U., Rabbitts T.H.;  
 RT "Comparison of the hinge-coding segments in human immunoglobulin gamma  
 RT heavy chain genes and the linkage of the gamma 2 and gamma 4 subclass  
 RT genes.";  
 RL EMBO J. 1:403-407(1982).  
 RN [4]  
 RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).  
 RX MEDLINE=81007873; PubMed=6774012;  
 RA Wang A.-C., Tung E., Fudenberg H.H.;  
 RT "The primary structure of a human IgG2 heavy chain: genetic,  
 RT evolutionary, and functional implications.";  
 RL J. Immunol. 125:1048-1054(1980).  
 RN [5]  
 RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).  
 RX MEDLINE=80001357; PubMed=113060;  
 RA Connell G.E., Parr D.M., Hofmann T.;  
 RT "The amino acid sequences of the three heavy chain constant region  
 RT domains of a human IgG2 myeloma protein.";  
 RL Can. J. Biochem. 57:758-767(1979).

[16] SEQUENCE OF 238-275 (ZIE).  
 RX MEDLINE=80114419; PubMed=118920;  
 RA Hofmann T., Parr D.M.;  
 RT "A note of the amino acid sequence of residues 381-391 of human  
 immunoglobulin gamma chains.";  
 RL Mol. Immunol. 16:923-925(1979).  
 [17]  
 RN REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).  
 RA Hofmann T., Parr D.M.;  
 RL Submitted (MAR-1980) to the PIR data bank.  
 [18]  
 RN SEQUENCE OF 1-121 (DOT).  
 RX MEDLINE=95255298; PubMed=7737190;  
 RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;  
 RT "Characterization of the two unique human anti-flavin monoclonal  
 immunoglobulins.";  
 RL Eur. J. Biochem. 228:886-893(1995).  
 [19]  
 RN DISULFIDE BONDS.  
 RX MEDLINE=72033500; PubMed=4940472;  
 RA Milstein C., Frangione B.;  
 RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";  
 RL Biochem. J. 121:217-225(1971).  
 [10]  
 RN DISULFIDE BONDS.  
 RX MEDLINE=69064124; PubMed=5782707;  
 RA Frangione B., Milstein C., Pink J.R.L.;  
 RT "Structural studies of immunoglobulin G.";  
 RL Nature 221:145-148(1969).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; J00230; AAB59393.1; -.  
 DR PIR; A93906; G2HU.  
 DR HSSP; P01857; 1FC1.  
 DR Genew; HGNC:5526; IGHG2.  
 DR MIM; 147110; -.  
 DR GO; GO:0005624; C:membrane fraction; NAS.  
 DR GO; GO:0003823; F:antigen binding activity; TAS.  
 DR GO; GO:0004955; P:immune response; NAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig-cl.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00407; IGcl; 2.  
 DR PROSITE; PS00835; IG-LIKE; 3.  
 DR PROSITE; PS00390; IG\_MHC; 2.  
 KW Immunoglobulin domain; Immunoglobulin C region.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 98 CH1.  
 FT DOMAIN 99 110 HINGE.  
 FT DOMAIN 111 219 CH2.  
 FT DOMAIN 220 326 CH3.  
 FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 27 83 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 140 200  
 FT DISULFID 246 304  
 FT SITE 156 156 AT OR NEAR THE COMPLEMENT-BINDING SITE.  
 FT MOD\_RES 326 326 REMOVED POST-TRANSLATIONALLY (PROBABLE).  
 FT VARIANT 60 60 S -> A (IN MYELOMA PROTEINS TIL & ZIE).  
 FT CONFLICT 109 109 /FTID=VAR\_003889.  
 C -> S (IN REF. 3).

SQ SEQUENCE 326 AA; 35884 MW; 8310878C6878CF9C CRC64;  
 Query Match 40.4%; Score 1148; DB 1; Length 326;  
 Best Local Similarity 72.2%; Pred. No. 8.6e-75;  
 Matches 226; Conservative 18; Mismatches 33; Indels 36; Gaps 6;  
 Qy 253 KAVTSNVASVQCDSNDKFP-----VYKYP-----GKGCPYL----- 283  
 Db 16 RSTSESTAALGCLVKDYFPEPTVSMNSGALTSVGVHFPVAVLQSSGLYSLSVTVVPSNN 75  
 Qy 284 -GDEGTDLDYYPEDTSGDKVRPH-----TPPCAPAPALGAPSVFLPPPKKDTLMIS 338  
 Db 76 FGTQTYTCNVDHKP-SNTKVDKTVKCKVCPCPCAPPVAG-PSVFLPPPKKDTLMIS 133  
 Qy 339 RTEVTCTVVVDVSHEDPEVKFNWYDGVGVHNAKTPREOVNSTYRVSVLTVLHQQDL 398  
 Db 134 RTEVTCTVVVDVSHEDPEVQFNWYDGVGVHNAKTPREOVNSTYRVSVLTVLHQQDL 193  
 Qy 399 NGKEYKCKVSNKALPVIETKISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYP 458  
 Db 194 NGKEYKCKVSNKGLPAPIETKISKAKGQPREPOVYTLPPSREEMTKNQVSLTCLVKGFYP 253  
 Qy 459 SDIAVWESNGQPNENYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHEALHN 518  
 Db 254 SDIAVWESNGQPNENYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHEALHN 313  
 Qy 519 HYTKQKSLSLSPGK 531  
 Db 314 HYTKQKSLSLSPGK 326  
 RESULT 4  
 GC4\_HUMAN  
 ID GC4\_HUMAN STANDARD; PRT; 327 AA.  
 AC P01861;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ig gamma-4 chain C region.  
 GN IGHG4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83157104; PubMed=6299662;  
 RA Ellison J.W., Buxbaum J.N., Hood L.E.;  
 RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";  
 RL DNA 1:11-18(1981).  
 RN [2]  
 RP SEQUENCE OF 1-30 AND 81-326.  
 RX MEDLINE=70207560; PubMed=4192699;  
 RA Pink J.R.L., Buttery S.H., de Vries G.M., Milstein C.;  
 RT "Human immunoglobulin subclasses. Partial amino acid sequence of the  
 constant region of a gamma 4 chain.";  
 RL Biochem. J. 117:33-47(1970).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; K01316; AAB59394.1; ALT\_INIT.  
 DR PIR; A90933; G4HU.  
 DR PDB; 1ADO; 16-SEP-98.  
 DR Genew; HGNC:5528; IGHG4.  
 DR MIM; 147130; -.  
 DR GO; GO:0005624; C:membrane fraction; NAS.  
 DR GO; GO:0003823; F:antigen binding activity; TAS.

```

DR GO: 0006955; P: immune response; NAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; IGcl; 2.
DR PROSITE: PS00835; IG_LIKE; 3.
DR PROSITE: PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region; 3D-structure.
FT NON_TER 1 1
FT DOMAIN 1 98 CH1.
FT DOMAIN 99 110 HINGE.
FT DOMAIN 111 220 CH2.
FT DOMAIN 221 327 CH3.
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 83
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 141 201
FT DISULFID 247 305
SQ SEQUENCE 327 AA; 35940 MW; 3EDBD811EF208E7A CRC64;

Query Match 40.0%; Score 1134.5; DB 1; Length 327;
Best Local Similarity 71.2%; Pred. No. 7.9e-74;
Matches 222; Conservative 17; Mismatches 40; Indels 33; Gaps 3;

QY 253 KAVTSNVASVOCDSNPKFP-----VVKYPG-----KGCP 282
DQ 16 RSTSESTAALGCLVKDYFPEPTVSNWNSGALTSGVHTFPAVLQSSGLYSLSSVTPPSS 75
QY 283 LGDEGDTLDYIYEEDETEGDKVRPH---TCPPCPAPEALGAPSVFLFPPKPKDTLMISR 339
DQ 76 LGRTKTYTCNVDRHKPSNTKVDKRVESKYGPCCPAPFLGPGSVFLFPPKPKDTLMISR 135
QY 340 TPVTCVVDVSHEDSEVKFNWYDGVGVHNAKTKPREQYNSTYRVSVLVTLVHODMUN 399
DQ 136 TPVTCVVDVSDQEDPEVQFNWYDGVGVHNAKTKPREQYNSTYRVSVLVTLVHODMUN 195
QY 400 GREYKCKVSKNALPVPTKTKISAKGPPREPQVYTLPPSREMTKNQVSLTCLVKGYFYS 459
DQ 196 GREYKCKVSKNGLPSSLEKTKISAKGPPREPQVYTLPPSREMTKNQVSLTCLVKGYFYS 255
QY 460 DIAVESNQCQPNNTYKTTTPVLVDSGSLYSLKLVDSKRWQGNVFCVSNVGHALNHH 519
DQ 256 DIAVESNQCQPNNTYKTTTPVLVDSGSLYSLKLVDSKRWQGNVFCVSNVGHALNHH 315
QY 520 YTKSLSLSGK 531
DQ 316 YTKSLSLSLGK 327

RESULT 5
GC3_HUMAN STANDARD; PRT; 290 AA.
AC P01860;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ig gamma-3 chain C region (Heavy chain disease protein) (HDC).
GN IGHG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE (DISEASE PROTEIN WIS).
RX MEDLINE=81021548; PubMed=6774747;
RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;
RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:
gamma 3 heavy-chain disease protein Wis.";
RL Biochemistry 19:4304-4308 (1980).
RN [2]
RP REVISIONS, TO 12-97 (PROTEIN WIS).

```

```

RX MEDLINE=77118561; PubMed=402363;
RA Michaelson T.E., Frangione B., Franklin E.C.;
RT "Primary structure of the 'hinge' region of human IgG3. Probable
quadruplication of a 15-amino acid residue basic unit.";
J. Biol. Chem. 252:883-889 (1977).
[3]
RN REVISIONS TO 59-289 (PROTEIN WIS/DISEASE PROTEIN WIS).
RX MEDLINE=77021516; PubMed=823945;
RA Wolfenstein-Rodei C., Frangione B., Prelli F., Franklin E.C.;
RT "The amino acid sequence of 'heavy chain disease' protein ZUC.
Structure of the PC fragment of immunoglobulin G3.";
Biochem. Biophys. Res. Commun. 71:907-914 (1976).
[4]
RN SEQUENCE FROM N.A. (DISEASE PROTEIN WIS).
RX MEDLINE=82247835; PubMed=6808505;
RA Alexander A., Steinmetz M., Barritault D., Frangione B.,
Franklin E.C., Hood L., Buxbaum J.N.;
RT "Gamma Heavy chain disease in man: cDNA sequence supports partial
gene deletion model.";
Proc. Natl. Acad. Sci. U.S.A. 79:3260-3264 (1982).
CC -1- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA
INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11
NORMALLY PRESENT IN THE HINGE REGION.
CC -1- MISCELLANEOUS: THE HEAVY CHAIN DISEASE PROTEIN WIS IS SHOWN.
CC -1- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE
REF.2.
CC -1- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION
AND ALL OF THE CH1 REGION.
CC -1- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL
OF THE CH1 REGION, AND PART OF THE HINGE COMPARED WITH NORMAL
GAMMA-3 HEAVY CHAINS.
CC -1- MISCELLANEOUS: DISEASE PROTEIN WIS MAY REPRESENT AN ALLELIC FORM
OR ANOTHER GAMMA CHAIN SUBCLASS.
CC -1- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR
TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE
IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE
SEGMENT (12-28).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; J00231; AAA52805.1; ALT_SEQ.
DR HSSP; P01857; IFC1.
DR Genew; HGNC:5527; IGHG3.
DR MIM; 147120; -.
DR GO: 0005624; C: membrane fraction; NAS.
DR GO: 0003823; F: antigen binding activity; TAS.
DR GO: 0006955; P: immune response; NAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig_C1.
DR Pfam: PF00047; Ig; 2.
DR SMART: SM00407; IGcl; 1.
DR PROSITE: PS00835; IG_LIKE; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat;
Pyridolone carboxylic acid.
FT DOMAIN 12 73 HINGE.
FT DOMAIN 74 183 CH2.
FT DOMAIN 184 289 CH3.
FT REPEAT 29 43
FT REPEAT 44 58
FT REPEAT 59 73
FT MOD_RES 1 1
FT CARBOHYD 6 6
FT DISULFID 7 7
FT DISULFID 24 24
FT DISULFID 27 27

```



```

FT DISULFID 33 33 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 39 39 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 42 42 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 48 48 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 54 54 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 57 57 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 63 63 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 69 69 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 72 72 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT CARBOHYD 140 140 N-LINKED (GLCNAC...).
FT MOD_RES 290 290 REMOVED POST-TRANSLATIONALLY.
FT VARIANT 126 127 QV -> EB (IN ZUC).
FT VARIANT 134 134 /FTID=VAR 003890.
FT VARIANT 139 139 P -> L (IN OMM).
FT VARIANT 182 182 /FTID=VAR 003891.
FT VARIANT 227 227 F -> Y (IN OMM).
FT VARIANT 227 227 /FTID=VAR 003892.
FT VARIANT 279 279 T -> A (IN OMM).
FT VARIANT 279 279 /FTID=VAR 003893.
FT VARIANT 279 279 S -> N (IN OMM).
FT VARIANT 279 279 /FTID=VAR 003894.
FT VARIANT 279 279 MISSING (IN ZUC).
FT VARIANT 279 279 /FTID=VAR 003895.
FT VARIANT 279 279 F -> Y (IN OMM).
FT VARIANT 279 279 /FTID=VAR 003896.
SQ SEQUENCE 290 AA; 32331 MW; E69CBC95705B2F46 CRC64;

Query Match 39.5%; Score 1121.5; DB 1; Length 290;
Best Local Similarity 78.7%; Pred. No. 5.7e-73;
Matches 211; Conservative 11; Mismatches 21; Indels 25; Gaps 2;

QY 264 CDSKDFPVKYGKCPGLTGDGDTLDYYPEDTEGDKVRHTCPCPAPEALGAPS 323
Db 48 CDTPPCPCPPEP-KSCDT-----PPPCPCPAPELLGGPS 82
QY 324 VLFPPKPKDLMISRTPEVTCVVDVSHDEDEKFNWYDGVVHNNAKTKPREQYNST 383
Db 83 VLFPPKPKDLMISRTPEVTCVVDVSHDEDEKFNWYDGVVHNNAKTKPREQYNST 142
QY 384 YRVSVLTVLHQDLNMGKCKYKCKSVKALPVETKISKAKGPPEQVYTLPPPREMT 443
Db 143 FRYVSVLTVLHQDLNMGKCKYKCKSVKALPAPIETKISKAKGPPEQVYTLPPPREMT 202
QY 444 KQVSLTCLVKGYFIPSDIAVESNCGPENNYKTPPVLDSDGSPFLYSKLVDSKRWQO 503
Db 203 KQVSLTCLVKGYFIPSDIAVESNCGPENNYKTPPVLDSDGSPFLYSKLVDSKRWQO 262
QY 504 GNVFSCVNHAEALHNNHYTKQSLSPGK 531
Db 263 GNVFSCVNHAEALHNNHYTKQSLSPGK 290

RESULT 6
GC_RABIT STANDARD; PRT; 323 AA.
AC P01870;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig gamma chain C region.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84030930; PubMed=6313520;
RA Bernstein K.E., Alexander C.B., Mage R.G.;
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant
RL F-1 haplotype";
RN Immunogenetics 18:387-397(1983).
RP SEQUENCE OF 1-128.

```

```

RX MEDLINE=76135469; PubMed=1243651;
RA Pratt D.M., Mole L.E.;
RT "Sequence studies on the constant region of the Fd sections of rabbit
RL immunoglobulin G of different allotype.";
RN Biochem. J. 151:337-349(1975).
RP [3]
RX SEQUENCE OF 88-266 FROM N.A.
RA MEDLINE=83299917; PubMed=6193512;
RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;
RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma
RL heavy chain and identification of two genomic C gamma genes.";
RN Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
RP [4]
RX SEQUENCE OF 132-161.
RA MEDLINE=70110015; PubMed=5461106;
RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;
RT "Sequence studies of the Fd section of the heavy chain of rabbit
RL immunoglobulin G.";
RN Biochem. J. 116:249-259(1970).
RP [5]
RX SEQUENCE OF 129-131 AND 155-322.
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;
RL (in) Killander J. (eds.);
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,
RL Stockholm (1967).
CC -1- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER,
CC 104-THR AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15
CC MARKERS AND REF.5 THE E15 MARKER.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M16426; AAA31289.1;
DR PIR; A91749; GHRE.
DR HSP; P01857; IPC1.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; ig; 2.
DR SMART; SM00407; Igcl; 2.
DR PROSITE; PS00835; IG_LIKS; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Repeat.
FT NON_TER 1 1
FT DOMAIN 6 96 IG-LIKE 1.
FT DOMAIN 114 213 IG-LIKE 2.
FT DOMAIN 222 318 IG-LIKE 3.
FT VARIANT 104 104 T -> M (IN D11 MARKER).
FT VARIANT 185 185 T -> A (IN E15 MARKER).
FT CONFLICT 48 48 N -> E (IN REF. 2).
FT CONFLICT 71 71 V -> VPV (IN REF. 2).
FT CONFLICT 144 144 Q -> B (IN REF. 3 AND 4).
FT CONFLICT 173 173 N -> D (IN REF. 5).
FT CONFLICT 187 187 Q -> E (IN REF. 5).
FT CONFLICT 201 201 N -> D (IN REF. 5).
FT CONFLICT 218 218 Q -> B (IN REF. 5).
FT CONFLICT 233 233 E -> Q (IN REF. 5).
FT CONFLICT 246 246 N -> D (IN REF. 5).
FT CONFLICT 256 256 E -> G (IN REF. 5).
FT CONFLICT 260 260 N -> D (IN REF. 5).
FT CONFLICT 266 266 E -> Q (IN REF. 5).
FT CONFLICT 280 280 Y -> W (IN REF. 5).
FT CONFLICT 284 284 N -> S (IN REF. 5).
SQ SEQUENCE 323 AA; 35404 MW; 69E8AA118D579A8B CRC64;

Query Match 32.4%; Score 921; DB 1; Length 323;
Best Local Similarity 61.1%; Pred. No. 1.3e-58;

```





CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; J00451; AAB59655.1; -;  
 DR EMBL; V01526; CAA24767.1; ALT\_SEQ.  
 DR PIR; A02156; G3MSM.  
 DR HSSP; P01857; 1FC1.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig cl.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00407; Igc1; 2.  
 DR PROSITE; PS00835; IG\_LIKE; 3.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 FT Transmembrane; Alternative splicing.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 113 HINGE.  
 FT DOMAIN 114 223 CH2.  
 FT DOMAIN 224 327 CH3.  
 FT TRANSMEM 346 362 POTENTIAL.  
 FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).  
 FT CONFLICT 333 333 E -> G (IN REF. 2).  
 FT CONFLICT 342 342 E -> Q (IN REF. 2).  
 FT CONFLICT 388 388 P -> F (IN REF. 2).  
 SQ SEQUENCE 398 AA; 43929 MW; CF7E264B50A41B95 CRC64;

Query Match 29.2%; Score 830; DB 1; Length 398;  
 Best Local Similarity 65.9%; Pred. No. 5.5e-52;  
 Matches 149; Conservative 32; Mismatches 43; Indels 2; Gaps 1;  
 QY 306 RPHTCP--PCPAPEALGAPSVLPFPKPKDTLMISRTPEVTVVDVSHEDPEVKFNWTV 363  
 DB 102 KSTPTGSSCPGNIIGGVSFVFPKPKDLMISLTPKTVVDVSDDDPDVHVSFV 161  
 QY 364 DGEVHNATKPREQNSTYRVSVLTVLHDPMLNGKEYCKVKSNKALPVIETISK 423  
 DB 162 DNKEVHTAWTPREAQVNSTFRVSVALPIQHODMGRKSEFKCKVNNKALPAPIERTISK 221  
 QY 424 KGQPREPOVTLPPREEMTKQVSLTCLVKGYPSPDIADVESNGQPNKYTPPVLD 483  
 DB 222 KGRAQTPQVYTIPTPREQMSKKVSUTCLVTNPFSEISVEMERNGELQDYKNTTPILD 281  
 QY 484 SDGSFFLYSKLTVDKSRQCGNVFSCVMHEALHNYHTOKSLSLSP 529  
 DB 282 SDGTFLYSLKLTVDTSKLGELFTCSVHVALHNNHTQNLRSRSP 327

RESULT 11  
 GC1\_MOUSE ID GC1\_MOUSE STANDARD; PRT; 324 AA.  
 AC P01568;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DE 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ig gamma-1 chain C region secreted form.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxId=10090;  
 RX SEQUENCE FROM N.A.  
 RX MEDLINE=60045036; PubMed=115593;  
 RA Honjo T.; Obata M.; Yamawaki-Kataoka Y.; Kataoka T.; Kawakami T.;  
 RA Takahashi N.; Mano Y.;  
 RA "Cloning and complete nucleotide sequence of mouse immunoglobulin  
 RT gamma 1 chain gene."  
 RL Cell 18:559-568(1979).

[2] SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).  
 RX MEDLINE=8020559; PubMed=6769752;  
 RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,  
 RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;  
 RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences  
 RT cloned in a bacterial plasmid."  
 RL Gene 9:87-97(1980).  
 [3] SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).  
 RX MEDLINE=80012837; PubMed=113776;  
 RA Rogers J., Clarke P., Salser W.;  
 RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin  
 RT heavy chain."  
 RL Nucleic Acids Res. 6:3305-3321(1979).  
 [4] SEQUENCE (MYELOMA PROTEIN MOPC 21).  
 RX MEDLINE=78242288; PubMed=98524;  
 RA Aducci K.;  
 RT "Evolution of immunoglobulin subclasses. Primary structure of a  
 RT murine myeloma gamma1 chain."  
 RL J. Biol. Chem. 253:6068-6075(1978).  
 [5] DISULFIDE BONDS (MOPC 21).  
 RX MEDLINE=73008889; PubMed=5073237;  
 RA Svasti J., Milstein C.;  
 RT "The disulfide bridges of a mouse immunoglobulin G1 protein."  
 RL Biochem. J. 126:837-850(1972).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Secreted;  
 CC IsoId=P01868-1; Sequence=Displayed;  
 CC Note=May be the major isoform;  
 CC Name=Membrane-bound;  
 CC IsoId=P01869-1; Sequence=External;  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; V00793; CAA24172.1; -;  
 DR EMBL; V00793; CAA24173.1; -;  
 DR EMBL; V00793; CAA24174.1; -;  
 DR EMBL; V00793; CAA24175.1; -;  
 DR EMBL; V00795; CAA24176.1; -;  
 DR PIR; A02159; G1MS.  
 DR PDB; 1IGC; 03-JUN-95.  
 DR GlycoSuiteDB; P01868; -;  
 DR MGD; MGI:96446; Igh-4.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig cl.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00407; Igc1; 2.  
 DR PROSITE; PS00835; IG\_LIKE; 3.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 KW Alternative splicing; 3D-structure.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 110 HINGE.  
 FT DOMAIN 111 217 CH2.  
 FT DOMAIN 218 324 CH3.  
 FT DISULFID 27 82  
 FT DISULFID 102 102  
 FT DISULFID 104 104  
 FT DISULFID 107 107  
 FT DISULFID 109 109  
 FT DISULFID 109 109 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).

```

FT DISULFID 138 198
FT CARBOHYD 174 174
FT FT N-LINKED (GLCNAC. . .)
FT /FTIG=CAR_000055.
FT DISULFID 244 302
FT MOD_RES 324 324
FT CONFLICT 276 276
FT CONFLICT 278 278
FT CONFLICT 278 278
SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;

Query Match
Best Local Similarity 29.1%; Score 826.5; DB 1; Length 324;
Matches 166; Conservative 50; Mismatches 68; Indels 75; Gaps 8;

QY 196 QENSLYT---IPKGF-----GSHLLPFAFLHGNPWLNCNCEILYFRRLWQ 237
D 18 QTNMVTGLCLVKGYFPEPVTVTVNSGSLSSGVTHTF-AVLQSD-----LY 62
QY 238 DNAENYVYKQVVDVAVNTSVASVQDNDKFPVYKYPGKCPGLDGGDTLDLYDYPE 297
D 63 TLSSSVTVFPSSPRPSETVTCNA-----HPA 88
QY 298 EDTEGDK--VRPHCTPP--CPAPEALGAPSVFLPPPKDTLMISRTPEVTCVVVDVSH 352
D 89 SSTKVDKIVPRDCCKPCICTVPE---VSSVFIAPPKPKDVLITLTPKVTVCVVVDISK 145
QY 353 EDPEVKFNWYDGVVHNAKTPREEQYNSTYRVVSVLTVLHQDLNGLKEYCKKVSNNAL 412
D 146 DQPEVQESMFVDDVEVHTAQTQPREQNFSTFRSVELPIHQDLNGLKEYCKKRVNSAF 205
QY 413 VPYEKTSKAKQDREPVQVTLPPSREMTNQVSLTCLVKGYFSPDAVWESNGOPE 472
D 206 PAPIEKTSKTKGRPKAPQVVTIPPPKQMAKDVKSLTCTMTDFPEDIIVVWONGOPA 265
QY 473 NNYKTPPVLDSDGSFFLYSLKTVDKSRWQGNVFCSCVWHEALHNYTKSLSPGK 531
D 266 ENYKTPQFMINGTSFYTSKLVNOKSNWAGNTFTCSVLHGLNHHHTKSLHSFGK 324

RESULT 12
GCC_RAT ID GCB_RAT STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Bruggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gere 74:473-482(1988).
CC -!- SIMILARITY: Contains 3 immunoglobulin-like domains.
DR PIR; PS0018; PS0018.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; ig; 3.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00835; IG-LIKE; 2.
DR PROSITE; PS00290; IG-LIKE; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Repeat.
FT NON_TER 1
FT DOMAIN 6 96 IG-LIKE 1.
FT DOMAIN 124 223 IG-LIKE 2.
FT DOMAIN 232 328 IG-LIKE 3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80

```

```

FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match
Best Local Similarity 29.1%; Score 826.5; DB 1; Length 333;
Matches 158; Conservative 34; Mismatches 63; Indels 27; Gaps 3;

QY 253 KAVTSNVA-SVQDNDKFPVYKYPGK--CPTLGDGEGDTLDLYYPEEDTEGDKVRPH 309
D 76 QTVTCNVAHPASSTKVDKVERNGGIGHKPTC-----PT 111
QY 310 CPPCPAPEALGAPSVFLPPPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVVH 369
D 112 CHKCPVPELLGGPSVFIFPPPKDILLISQNAKVTCTVVVDVSEEPDQVQFVFNVEVH 171
QY 370 NAKTKPREQYNSTYRVVSVLTVLHQDLNGLKEYCKKVSNNALPVPYEKTSKAKGPRE 429
D 172 TAQTQPREQYNSTFRVVSALPIQHDMMSGKEFKCKVNNKALPSPYEKTSKPKGLVRK 231
QY 430 PQVYTLPPSREMTNQVSLTCLVKGYFSPDAVWESNGOPENNYKTPPVLDSDGSFF 489
D 232 PQVYVNGPTEQLTEQTVSLTCLTSGFLPNDIGVWTSNGHIKQVKNTEPVWDSGDSFF 291
QY 490 LYSKLTVDKSRWQGNVFCSCVWHEALHNYTKSLSPGK 531
D 292 MYSKLVNRSRWSRAPPVFCVSVHGLNHHVKSISRPFGK 333

RESULT 13
GCC_RAT ID GCB_RAT STANDARD; PRT; 329 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2C chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166903; PubMed=3127222;
RA Bruggemann M.; Delmaestro-Galfre P.; Waldmann H.; Calabi P.;
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
region cDNA: extensive homology to mouse gamma 3.";
RL Eur. J. Immunol. 18:317-319(1988).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; X07189; CAA30169.1; -.
DR PIR; S00847; S00847.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; ig; 2.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00835; IG-LIKE; 3.
DR PROSITE; PS00290; IG-LIKE; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1

```

```
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 222 CH2.
FT DOMAIN 223 329 CH3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 143 203 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 249 307 INTERCHAIN (WITH A HEAVY CHAIN).
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 28.9%; Score 821.5; DB 1; Length 329;
Best Local Similarity 64.2%; Pred. No. 1.7e-51;
Matches 147; Conservative 38; Mismatches 41; Indels 3; Gaps 1;

QY 306 RPHTCPP---CPAPEALGAPSVLFPFKPKDTLMISRTPEVTCVVVDVSHEDPEVKFMY 362
DB 101 RPKRPPTDICSDDNLGRPSVFIFPKPKDILMITLTPKVTCCVVVDVSEEEFDVQFSWF 160
QY 363 VQGVVHNAKTPREYQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTSK 422
DB 161 VQNRVFTAQDPEHQLNGTRFVWSTLTIQHQMMSGKEFKCKVNNKDLPSPIETSK 220
QY 423 AKQGPPEQVYTLPPSREEMTKNQVSLTCLLVKGFPYPSDIAVEWESNGQPENNYKTPPV 482
DB 221 PRGKARTPVYTIIPPREQMSKKNKVSITCMVTSFYFASISVEWERNGELEQDYKNTLP 280
QY 483 DSDGSFSLYSKLTVDKSRVQQGNFVSCSVNHEALHNYHTQKLSLSPGK 531
DB 281 DSDSEYFLYSKLSVDYTNWGRDIYTCVVVHEALHNNHTQKNLSRSPGK 329

RESULT 14
GC1M_MOUSE STANDARD; PRT; 393 AA.
AC P01869;
DT 21-JUL-1986 (Rel. 0.; Created)
DT 01-AUG-1991 (Rel. 19; Last sequence update)
DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE Ig gamma-1 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T.; Obata M.; Yamawaki-Kataoka Y.; Kataoka T.; Kawakami T.;
RA Takahashi N.; Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene."
RL Cell 18:559-568 (1979).
RN [2]
RP SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M.; Cowman A.F.; Gerondakis S.D.; Adams J.M.; Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
conserved transmembrane sequence and a 28-residue intracellular
domain."
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012 (1982).
RN [3]
RP SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J.; Choi R.; Souza L.; Carter C.; Word C.J.; Kuehl M.;
RA Eisenberg D.; Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
immunoglobulin gamma chains."
RL Cell 26:19-27 (1981).
RN [4]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y.; Nakai S.; Miyata T.; Honjo T.;
```

```
RT "Nucleotide sequences of gene segments encoding membrane domains of
immunoglobulin gamma chains."; Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627 (1982).
RL -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Membrane-bound;
CC IsoId=P01869-1; Sequence=Displayed;
CC Name=Secreted;
CC IsoId=P01868-1; Sequence=External;
CC Note=May be the major isoform;
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DB EMBL; V00793; CAA24172.1; -
DB EMBL; V00793; CAA24173.1; -
DB EMBL; V00793; CAA24174.1; -
DB PIR; B02159; G1MSM.
DB PDB; 15C8; 23-MAR-99.
DB PDB; 1AE6; 18-MAR-98.
DB PDB; 1CL7; 12-JAN-00.
DB PDB; 1FL1; 06-FEB-01.
DB PDB; 1F58; 29-DEC-99.
DB PDB; 1KCS; 24-JUL-02.
DB PDB; 1KCR; 11-MAY-02.
DB PDB; 25C8; 09-JUL-99.
DB MGD; MGI:96446; Igh-4.
DB InterPro; IPR007110; Ig-like.
DB InterPro; IPR003597; Ig_C1.
DB InterPro; IPR003006; Ig_MHC.
DB Pfam; PF00047; Ig; 3.
DB SMART; SM00407; IGc1; 2.
DB PROSITE; PS00835; IG_LIKE; 3.
DB PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane; 3D-structure.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198 N-LINKED (GLCNAC. . .).
FT CARBOHYD 174 174 POTENTIAL.
FT DISULFID 244 302 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 340 357 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 358 393
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;

Query Match 28.9%; Score 821.5; DB 1; Length 393;
Best Local Similarity 46.1%; Pred. No. 2.2e-51;
Matches 165; Conservative 50; Mismatches 68; Indels 75; Gaps 8;

QY 196 QENSLYT---IPKGF-----GSHLLPFAFLHGNPMLCNCEILYFRWLQ 237
DB 18 QTNSMVTGLCLVKGYFPEPVTVTVNSGSLSSGVHTFP-AVLQSD-----LY 62
QY 238 DNAENVVVMQVVDVKAVTSNVALSVQCDNSDKFPVVKYKPGCGFTLGDGDTLDLYYPE 297
DB 63 TLSSSVTVPSRPRSETVTCNVA-----HPA 88
QY 298 EDTGDK---VRPHTCPP---CPAPEALGAPSVLFPFKPKDTLMISRTPEVTCVVVDVSH 352
DB 89 SSTVKDKKIIVPRDCGCKPCICTYPE---VSSVFIPFPKPKDVLITLTPKVTCTVDISK 145
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:16:09 ; Search time 55.8167 Seconds  
(without alignments)  
2454.925 Million cell updates/sec

Title: US-10-068-426-5

Perfect score: 2839

Sequence: 1 MPELLLLLLPSLPHPIC.....MHEALHNYTKSLSLSPGK 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1616.5	56.9	626	4 Q8NG39	Q8ng39 homo sapien
2	1615.5	56.9	626	4 Q8NLF3	Q8nlf3 homo sapien
3	1208	42.6	701	4 Q96PQ8	Q96pq8 homo sapien
4	1205	42.4	471	4 Q8TC77	Q8tc77 homo sapien
5	1141	40.2	521	4 Q8N4Y9	Q8n4y9 homo sapien
6	1135	40.0	473	4 Q8TC63	Q8tc63 homo sapien
7	1130	39.8	509	4 Q8NF17	Q8nf17 homo sapien
8	1032	36.4	734	11 Q35930	Q35930 mus musculu
9	927	32.7	677	6 Q28256	Q28256 canis famli
10	891	31.4	337	6 Q95M34	Q95m34 equus cabal
11	842.5	29.7	469	11 Q8R3V9	Q8r3v9 mus musculu
12	839.5	29.6	463	11 Q9SLC4	Q9slc4 mus musculu
13	835.5	29.4	437	11 Q9RLA4	Q9rla4 mus musculu
14	814.5	28.7	473	11 Q9D8L4	Q9d8l4 mus musculu
15	805	28.4	468	11 Q9SL31	Q9sl31 mus musculu
16	805	28.4	473	11 Q99L25	Q99l25 mus musculu

17	771	27.2	473	11 Q91Z05	Q91z05 mus musculu
18	771	27.2	474	11 Q8R3H6	Q8r3h6 mus musculu
19	359.5	12.7	375	4 Q9BSZ1	Q9bsz1 homo sapien
20	359.5	12.7	597	4 Q9BU10	Q9bu10 homo sapien
21	359.5	12.7	597	4 Q9BQ88	Q9bq88 homo sapien
22	359.5	12.7	597	4 Q9BBB9	Q9bbb9 homo sapien
23	353	12.4	614	4 Q96GA6	Q96ga6 homo sapien
24	349.5	12.3	588	4 Q8WUX4	Q8wux4 homo sapien
25	349.5	12.3	613	4 Q96EY0	Q96ey0 homo sapien
26	349.5	12.3	613	4 Q8WUK1	Q8wuk1 homo sapien
27	349.5	12.3	618	4 Q96AA6	Q96aa6 homo sapien
28	342	12.0	613	11 Q8VCX7	Q8vcx7 mus musculu
29	302.5	10.7	342	11 Q91XL1	Q91xl1 mus musculu
30	301	10.6	347	4 Q8N4F5	Q8n4f5 homo sapien
31	296	10.4	1328	5 Q21043	Q21043 caenorhabdi
32	284.5	10.0	486	11 Q91Z07	Q91z07 mus musculu
33	282.5	10.0	384	4 Q9UP60	Q9up60 homo sapien
34	282.5	10.0	493	4 Q8NCL6	Q8nc16 homo sapien
35	282.5	10.0	494	4 Q96K68	Q96k68 homo sapien
36	282.5	10.0	496	4 Q96KX8	Q96kx8 homo sapien
37	279.5	9.8	496	4 Q96DK0	Q96dk0 homo sapien
38	279.5	9.8	499	4 Q8NSK4	Q8nsk4 homo sapien
39	277.5	9.8	497	4 Q8WY24	Q8wy24 homo sapien
40	277.5	9.8	500	4 Q9BRV0	Q9brv0 homo sapien
41	276	9.7	557	4 Q8NSV4	Q8nsv4 homo sapien
42	274	9.7	416	4 Q9NPP6	Q9npp6 homo sapien
43	274	9.7	426	11 Q9DCD9	Q9dcd9 mus musculu
44	273	9.6	487	11 Q99KA4	Q99ka4 mus musculu
45	271.5	9.6	1527	5 Q9VZZ4	Q9vzz4 drosophila

ALIGNMENTS

RESULT 1

Q8NG39 ID Q8NG39 PRELIMINARY; PRT; 626 AA.  
AC Q8NG39;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Platelet glycoprotein Ib alpha.  
GN GPIBA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Matsubara Y., Murata M., Sugita K., Ikeda Y.;  
RT "Identification of a novel point mutation in platelet glycoprotein  
RT Iba, Gly to Ser at residue 233, in a Japanese family with platelet-  
RT type von Willebrand diseasee.";  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB086948; BAC10305.1; -;  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR000483; LRR\_Cterm.  
DR InterPro; IPR000372; LRR\_Nterm.  
DR InterPro; IPR003591; LRR\_typ.  
DR InterPro; IPR002965; P\_rich\_extensn.  
DR Pfam; PF00560; LRR; 1.  
DR Pfam; PF01463; LRRCT; 1.  
DR Pfam; PF01462; LRRNT; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR PRINTS; PR03217; PRICHEXTENS.  
DR SMART; SM00082; LRRCT; 1.  
DR SMART; SM00013; LRRNT; 1.  
DR SMART; SM00369; LRR\_TYP; 5.  
DR PROSITE; PS05056; LRR\_TYPICAL; 1.  
SQ SEQUENCE 626 AA; 69985 MW; C7931FD07458B17F CRC64;

Query Match 56.9%; Score 1616.5; DB 4; Length 626;  
Best Local Similarity 84.2%; Pred. No. 2.1e-125;



Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;

Qy 1 MLLLLLLLLLPSPLPHPICEVSKVASHLEVNCNRLTALPPDLKDTTILHLSENLLY 60  
 Db 1 MLLLLLLLLLPSPLPHPICEVSKVASHLEVNCNRLTALPPDLKDTTILHLSENLLY 60

Qy 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLPLVGLTDLSHQLOSLPLGQTLPALTV 120  
 Db 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLPLVGLTDLSHQLOSLPLGQTLPALTV 120

Qy 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180  
 Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180

Qy 181 AGLNGLENLDTLLQNSLYTIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRFWLQDNA 240  
 Db 181 AGLNGLENLDTLLQNSLYTIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRFWLQDNA 240

Qy 241 ENVVWKQVVDVKAVTSNVASVQCDNSDKFPVVKYPKGCGPTLGDGDTLDLYDYYPEEDT 300  
 Db 241 ENVVWKQVVDVKAVTSNVASVQCDNSDKFPVVKYPKGCGPTLGDGDTLDLYDYYPEEDT 300

Qy 301 EGDKVR-----PHTCP-----PCPA---PEALGAPSVFLFPK-- 330  
 Db 301 EGDKVRATRTVVKFPPTKAHTTPWGLFYSWSTASLDSQMPSSLHPTQBSTKEQTTFPPRWT 360

Qy 331 PKDTL-----MISRTPEVT 344  
 Db 361 PNFTLHMSITFSKTPKST 379

## RESULT 2

Q8N1F3 PRELIMINARY; PRT; 626 AA.

ID Q8N1F3  
 AC Q8N1F3;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Glycoprotein Ib (Platelet), alpha polypeptide.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RA Strausberg R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC027955; AAH27955.1; -.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000483; LRR\_Cterm.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR InterPro; IPR003591; LRR\_tyr.  
 DR InterPro; IPR002965; P\_Rich\_extensn.  
 DR Pfam; PF00560; LRR; 1.  
 DR Pfam; PF01463; LRRCT; 1.  
 DR Pfam; PF01462; LRRNT; 1.  
 DR PRINTS; PR00019; LEURICHRPT.  
 DR PRINTS; PR001217; PRICHEXTENS.  
 DR SMART; SM00082; LRRCT; 1.  
 DR SMART; SM00013; LRRNT; 1.  
 DR SMART; SM00369; LRR\_TYP; 5.  
 DR PROSITE; PS05056; LRR\_TYPICAL; 1.  
 SQ SEQUENCE 626 AA; 68969 MW; 1945FE52545DF573 CRC64;

Query Match 56.9%; Score 1615.5; DB 4; Length 626;  
 Best Local Similarity 84.2%; Pred. NO. 2.5e-125;  
 Matches 319; Conservative 5; Mismatches 20; Indels 35; Gaps 5;

Qy 1 MLLLLLLLLLPSPLPHPICEVSKVASHLEVNCNRLTALPPDLKDTTILHLSENLLY 60  
 Db 1 MLLLLLLLLLPSPLPHPICEVSKVASHLEVNCNRLTALPPDLKDTTILHLSENLLY 60

Qy 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLPLVGLTDLSHQLOSLPLGQTLPALTV 120  
 Db 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLPLVGLTDLSHQLOSLPLGQTLPALTV 120

Qy 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180  
 Db 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180

Qy 181 AGLNGLENLDTLLQNSLYTIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRFWLQDNA 240  
 Db 181 AGLNGLENLDTLLQNSLYTIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRFWLQDNA 240

Qy 241 ENVVWKQVVDVKAVTSNVASVQCDNSDKFPVVKYPKGCGPTLGDGDTLDLYDYYPEEDT 300  
 Db 241 ENVVWKQVVDVKAVTSNVASVQCDNSDKFPVVKYPKGCGPTLGDGDTLDLYDYYPEEDT 300

Qy 301 EGDKVR-----PHTCP-----PCPA---PEALGAPSVFLFPK-- 330  
 Db 301 EGDKVRATRTVVKFPPTKAHTTPWGLFYSWSTASLDSQMPSSLHPTQBSTKEQTTFPPRWT 360

Qy 331 PKDTL-----MISRTPEVT 344  
 Db 361 PNFTLHMSITFSKTPKST 379

RESULT 3

Q96P08 PRELIMINARY; PRT; 701 AA.

ID Q96P08  
 AC Q96P08;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Factor VII active site mutant immunoconjugate.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21477448; PubMed=11593034;  
 RA Hu Z.; Garen A.;  
 RT "Targeting tissue factor on tumor vascular endothelial cells and tumor  
 RT cells for immunotherapy in mouse models of prostatic cancer.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).  
 CC !- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 DR EMBL; AF272774; AAK58686.1; -.  
 DR HSPF; P00761; IAN1.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR000742; EGF\_2.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR001438; EGF\_II.  
 DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR002383; GLA\_blood.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003597; IG\_c1.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR InterPro; IPR000294; VitK\_dep\_GLA.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF00594; gla; 1.  
 DR Pfam; PF00047; ig; 2.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00010; EGFBLOOD.  
 DR PRINTS; PR00001; GLABLOOD.  
 DR SMART; SM00179; EGF\_CA; 1.  
 DR SMART; SM00069; GLA; 1.  
 DR SMART; SM00407; IGC1; 1.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.

Db	213	PSSSLGTQTYICNVNHHKPSNTKVDKVBPKSCDKTHTCPCPAPELGGPSVFLFPKPK	272
Qy	333	DTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQNSTYRVVSVLTV	392
Db	273	DTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQNSTYRVVSVLTV	332
Qy	393	LHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPPQVYTLPPSREEMTKNQVSLTCL	452
Db	333	LHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPPQVYTLPPSREDELTKNQVSLTCL	392
Qy	453	VKGFPYSDIAVEHESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFC	512
Db	393	VKGFPYSDIAVEHESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFC	452
Qy	513	HEALHNHYTOKLSLSFGK	531
Db	453	HEALHNHYTOKLSLSFGK	471
RESULT 5			
Q8N4Y9 PRELIMINARY; PRT; 521 AA.			
AC	Q8N4Y9	PRELIMINARY; PRT; 521 AA.	
DT	01-OCT-2002 (TrEMBLrel. 22, Created)		
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Hypothetical protein.		
OS	Homo sapiens (human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
NCBI_TaxID=9606;			
RN	[1]		
RS	SEQUENCE FROM N.A.		
RP	TISSUE=Primary B-Cells from Tonsils;		
RC	Strausberg R.;		
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC033178; AAH33178.1; -		
DR	InterPro; IPR007110; Ig-like.		
DR	InterPro; IPR003197; Ig cl.		
DR	InterPro; IPR003006; IG_MHC.		
DR	InterPro; IPR003196; Ig_v.		
DR	Pfam; PF00047; Ig; 4.		
DR	SMART; SM00407; IGcl; 3.		
DR	SMART; SM00406; IGV; 1.		
DR	PROSITE; PS50835; IG_LIKS; 4.		
DR	PROSITE; PS00290; IG_MHC; 2.		
KW	Hypothetical protein.		
SK	SEQUENCE 521 AA; 57156 MW; 2AC7D22872D6CAA2 CRC64;		
Query Match 40.2%; Score 1141; DB 4; Length 521;			
Best Local Similarity 64.1%; Pred. No. 4.5e-86;			
Matches 229; Conservative 18; Mismatches 56; Indels 54; Gaps 7;			
Qy	194	LLQENSLY-----TTPKGFPGSHLLPFPFLHGNPMLCNCEILYFRRLWLDQNAENVVYWK	247
Db	200	LLQENSLY-----TTPKGFPGSHLLPFPFLHGNPMLCNCEILYFRRLWLDQNAENVVYWK	240
Qy	248	QVVDVKAIVTSNVAIV-----QCDNSDKFPYKYPGK-----GGPTLGDGDTLDYD	294
Db	241	R-VELATPLGLDTHTCPCPCPEPKSCDTPPCPCPEPKSCDTPPCPCPEPKSCD-----	295
Qy	295	YPBEDTEGDKVRPHTPCPAPALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHED	354
Db	296	YPBEDTEGDKVRPHTPCPAPALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHED	344
Qy	355	PEVKFNWYVDGVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPV	414
Db	345	PEVQFKWYVDGVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPA	404
Qy	415	PIEKTISKAKGQPREPPQVYTLPPSREEMTKNQVSLTCLVKGFPSDIAVEHESNGQPEN	474
Db	405	PIEKTISKAKGQPREPPQVYTLPPSREEMTKNQVSLTCLVKGFPSDIAVEHESNGQPEN	464





RT "Organization of the equine immunoglobulin heavy chain constant region  
 RL genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";  
 RL Immunoblotology 199:105-119(1998).

DR EMBL: AJ300675; CAC44624.1; -  
 DR InterPro: IPR007110; Ig\_LIKE.  
 DR InterPro: IPR003597; Ig\_c1.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR Pfam: PF00047; Ig; 2.  
 DR SMART: SM00407; IgC1; 2.  
 DR PROSITE; PS00835; IG\_LIKE; 3.  
 DR PROSITE; PS00290; IG\_MHC; 2.  
 FT NON TER 1  
 SQ SEQUENCE 337 AA: 37438 MW: A608F2B01DEFDIF6 CRC64;

Query Match 31.48; Score 891; DB 6; Length 337;  
 Best Local Similarity 70.11; Pred. No. 1.3e-65;  
 Matches 157; Conservative 34; Mismatches 31; Indels 2; Gaps 1;

QY 310 CPPEALGAPSVFLPPPKDGLTLMISRTPEVTCVVDVSHEDPEVFNWYVDGVEVH 369  
 DB 114 CPKPAPELLGGPSVFIFPNPKDGLTMTPEVTCVVDVSGENPDVKNWYMGVEYR 173  
 QY 370 NAKTKPREQYNSTYRVSVLTVLHODWLNKGYKCKVSKNKAALPVPIKTIKAKQPRE 429  
 DB 174 TATTPRKEQFNSTYRVSVLRIQHODWLSGKFKCKVANNQALPOPIERTITTKGRSQE 233  
 QY 430 POYVTLPPSREMTKNOVSLTCLVKGFYPSDIAVEVESNGQP--ENNYKTTTPVLDSGS 487  
 DB 234 POYVTLAPHEDELKSKSVSLTCLVKGFYPSDIAVEVESNGQPPELEYKSTTQAQSDSGS 293  
 QY 488 FFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 531  
 DB 294 YFLYSKLSVDNRWQQTFTTCGVMHEALHNHYTKQNVSKNPGK 337

## RESULT 11

Q8R3V9 PRELIMINARY; PRT; 469 AA.  
 ID Q8R3V9  
 AC Q8R3V9;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical 52.0 kDa protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC024405; AAH24405.1; -  
 DR InterPro: IPR007110; Ig\_LIKE.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig; 4.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS00835; IG\_LIKE; 4.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 469 AA: 51976 MW: 534793F15D05457 CRC64;

Query Match 29.78; Score 842.5; DB 11; Length 469;  
 Best Local Similarity 47.18; Pred. No. 2.2e-61;  
 Matches 169; Conservative 47; Mismatches 68; Indels 75; Gaps 8;

QY 196 QENSLEYT---IPKGF-----GSHLLPFAFHGNPWLNCIELYFRRLQ 237  
 DB 163 QTNSSVTLGLVKGYFPEPVTVTNSSGSSGVHTFP-AVLQSD-----LY 207  
 QY 238 DNAENVVYKQVVDVKAITSNVASVQCDNSDKFPVKYKPGCGPTLGDGDTLDLYYPE 297  
 DB 208 TLSSSVTPSPSTWPSQTVCNVA-----HPA 227

QY 298 EDTSGDK---VRPHTCPP--CPAPEALGAPSVFLPPPKDGLTLMISRTPEVTCVVDVSH 352  
 DB 234 SSTVKDKKIVPRDCGCKPCICTVPE---VSSVFIFPPPKDGLTITLTPKVTCCVVDISK 290  
 QY 353 EDPEVKFNWYVDGVEVHNNAKTKPREQYNSTYRVSVLTVLHODWLNKGYKCKVSKNKAAL 412  
 DB 291 DDPEVQFSWFVDVVEVHTAQTCKPREQFNSTFRSVSELPIMHQDWLNKGFCKRVNSAAF 350  
 QY 413 PVPIKTIKAKQPREPOVYTLPPSREMTKNOVSLTCLVKGFYPSDIAVEVESNGQPE 472  
 DB 351 PAPIKTIKTKGRPKAPQVYTIIPPKKEQMAKQVSLTCHITDFFPEDITVEWQWNGQPA 410  
 QY 473 NNYKTTTPVLDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 531  
 DB 411 ENYKNTQPIMDTGSYFVYSLKLVQKSNWEAGNTFTCSVLHGLHNHHTKSLSHSPGK 469

## RESULT 12

Q99LC4 PRELIMINARY; PRT; 463 AA.  
 ID Q99LC4  
 AC Q99LC4;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Similar to RIKEN cDNA 1810060009 gene.  
 GN IGH-4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC003435; AAH03435.1; -  
 DR HSSP; P01842; 7FAB.  
 DR MGD; MGI:96446; Igh-4.  
 DR InterPro: IPR007110; Ig\_LIKE.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam; PF00047; Ig; 4.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS00835; IG\_LIKE; 4.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 SQ SEQUENCE 463 AA: 51007 MW: EAA674C6BBC30783 CRC64;

Query Match 29.61; Score 839.5; DB 11; Length 463;  
 Best Local Similarity 46.81; Pred. No. 3.9e-61;  
 Matches 168; Conservative 48; Mismatches 68; Indels 75; Gaps 8;

QY 196 QENSLEYT---IPKGF-----GSHLLPFAFHGNPWLNCIELYFRRLQ 237  
 DB 157 QTNSSVTLGLVKGYFPEPVTVTNSSGSSGVHTFP-AVLQSD-----LY 201  
 QY 238 DNAENVVYKQVVDVKAITSNVASVQCDNSDKFPVKYKPGCGPTLGDGDTLDLYYPE 297  
 DB 202 TLSSSVTPSPSTWPSQTVCNVA-----HPA 227  
 QY 298 EDTSGDK---VRPHTCPP--CPAPEALGAPSVFLPPPKDGLTLMISRTPEVTCVVDVSH 352  
 DB 228 SSTVKDKKIVPRDCGCKPCICTVPE---VSSVFIFPPPKDGLTITLTPKVTCCVVDISK 284  
 QY 353 EDPEVKFNWYVDGVEVHNNAKTKPREQYNSTYRVSVLTVLHODWLNKGYKCKVSKNKAAL 412  
 DB 285 DDPEVQFSWFVDVVEVHTAQTCKPREQFNSTFRSVSELPIMHQDWLNKGFCKRVNSAAF 344  
 QY 413 PVPIKTIKAKQPREPOVYTLPPSREMTKNOVSLTCLVKGFYPSDIAVEVESNGQPE 472  
 DB 345 PAPIKTIKTKGRPKAPQVYTIIPPKKEQMAKQVSLTCHITDFFPEDITVEWQWNGQPA 404  
 QY 473 NNYKTTTPVLDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 531

Db 405 ENYKNTQPMIDTSGSYFYSLKLVQKSNWEAGNTFTCSVLHGLNHHHTKSLSHSPGK 463

RESULT 13

Q9D8L4 PRELIMINARY; PRT; 437 AA.

AC Q9D8L4; 437

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Gamal heavy chain of Mab7 (Fragment).

GN IGh-4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI\_TaxID=10090;

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;

RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal antibody (Mab 7, its light and heavy chains) and construction of a single chain antibody (scFv).";

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF152372; AD40243.1; -.

DR HSSP: P01842; 7FAB.

DR MGD: MGI:96446; Igh-4.

DR InterPro: IPR007110; Ig-Like.

DR InterPro: IPR003006; IG\_MHC.

DR InterPro: IPR003596; IG\_V.

DR Pfam: PF00047; Ig; 4.

DR SMART: SM00406; IGV; 1.

DR PROSITE: PS00835; IG\_LIKE; 4.

DR PROSITE: PS00290; IG\_MHC; 1.

FT NON\_TER 1

FT NON\_TER 437

SQ SEQUENCE 437 AA; 48142 MW; 5C3A7B3BE7D697C CRC64;

Query Match 29.4%; Score 835.5; DB 11; Length 437;

Best Local Similarity 46.5%; Pred. No. 7.7e-61;

Matches 167; Conservative 49; Mismatches 68; Indels 75; Gaps 8;

QY 196 QENSLYV---IPKGF-----GSHLLPFAFLHGNPMLCNCEILYFRWLQ 237

DB 131 QTNMVTLAGLVKGYFPEFVTWNSGSLSSGVHTFP-AVLQSD-----LY 175

QY 238 DNAENYVVKQVVDVKAFTSNVASVQCDNSDKFPVYKPGKPTLGDSGDTLDYDYPE 297

DB 176 TLSSSVTVSSWSPSTVTCNVA-----HPA 201

QY 298 EDTGDK---VRPHTCPP--CPAPEALGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSH 352

DB 202 SSTKVDKIVPRDCGCKPCICTVPE--VSSVFIPPKPKDVLITLTPKVCVVVDISK 258

QY 353 EDPEVKFNKYVDGVEVHNKATKPREQYNSTRVSVLTVLHODWLNKGYCKYKSNKAL 412

DB 259 DDEVFQFSFVDDVEVHTAQTRREQFNSTRSVSELPIMHQDMLNKGKFKCRVNSAAF 318

QY 413 PVIPIKTSKAKQPREPOVYTLPPREEMTKNQVSLTCLVKGFYPSDIAVEWESNGOPE 472

DB 319 PAPIETIKTKGRPAQOVYTIIPPKQMAKDKVSLTCHITDFFPEDITVQWNGQDA 378

QY 473 NNYKTTPLVDSGSPFLYSKLTVDKSRWQOQNVFSCSVHGLNHHHTYTKSLSPGK 531

DB 379 ENYKNTQPMIDTSGSYFYSLKLVQKSNWEAGNTFTCSVLHGLNHHHTKSLSHSPGK 437

RESULT 14

Q9D8L4 PRELIMINARY; PRT; 473 AA.

AC Q9D8L4; 473

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE 181006009Rik protein.

GN IGh-1 OR 181006009RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=C57BL/6J; TISSUE=Pancreas;

RC MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Cabavant T.,

RA Fleischmann W., Gaasterland T., Giasi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schirimi L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barah G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., King B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,

RA Hayashizaki Y.

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EXBL: AK007918; BAB25349.1; -.

DR HSSP: P01842; 7FAB.

DR MGD: MGI:96443; Igh-1.

DR InterPro: IPR007110; Ig-Like.

DR InterPro: IPR003006; IG\_MHC.

DR InterPro: IPR003596; IG\_V.

DR Pfam: PF00047; Ig; 4.

DR SMART: SM00406; IGV; 1.

DR PROSITE: PS00835; IG\_LIKE; 4.

DR PROSITE: PS00290; IG\_MHC; 1.

SQ SEQUENCE 473 AA; 51699 MW; 9DE057A514475FBB CRC64;

Query Match 28.7%; Score 814.5; DB 11; Length 473;

Best Local Similarity 46.3%; Pred. No. 4.8e-59;

Matches 164; Conservative 47; Mismatches 80; Indels 63; Gaps 6;

QY 199 SLYTIPIKGF-----GSHLLPFAFLHGNPMLCNCEILYFRWLQDAENV 243

DB 162 TLGLVKVGYFPEFVTWNSGSLSSGVHTFPALLQSG-----LYTLSSV 206

QY 244 YVKKQVVDVKAFTSNVA-----SVQCDN--SDKFPVYKPGKCGPTLGDEGDTLDYDYPE 297

DB 207 TVTSNTWPSOTITCNVAHPASSTKVKKIEPRVITQNP---CPPLKE----- 251

QY 298 EDTGDKVRPHTCPPCPAPEALGAPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEV 357

DB 252 -----CPPCAAPDLGGPSVFIPPKIKDVLMLSLSPMVTCCVVVDVSEDDPDV 299

QY 358 KFNWYDGVGVHNKATKPREQYNSTRVSVLTVLHODWLNKGYCKYKSNKALPVPIE 417

DB 300 QISMFVNNVNEVHTAQTRREQFNSTRVSVLSALPIHQDWMGSKGFKCKYNNALPSPIE 359

QY 418 KTSKAKQPREPOVYTLPPREEMTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKT 477

DB 360 KTSKPRGPRAPQVYVLPPEAEEMTKKESLTCMTGTGFLPAEIAVDVTSNGTEQNYKN 419

QY 478 TTPVLDSDGSPFLYSKLTVDKSRWQOQNVFSCSVHGLNHHHTYTKSLSPGK 531

DB 420 TATVLDSDGSPFYVSKLVKQKSTWERSGLFACSVVHGLNHHHTTKTISRSLCK 473

RESULT 15

Q99L31

```
ID Q99L31 PRELIMINARY; PRT; 468 AA.
AC Q99L31;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Similar to RIKEN CDNA 1810060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003878; AA03878.1; -.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 1.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 468 AA; 51661 MW; 96352328B3332ADB CRC64;

Query Match      28.4%; Score 805; DB 11; Length 468;
Beet Local Similarity 60.7%; Pred. No. 2.9e-58;
Matches 150; Conservative 32; Mismatches 53; Indels 12; Gaps 3;

QY 295 YPEEDTEGDK-----VRPHTCPP--CPAPEALGAPSVFLFPKPKDITMISRTPEVT 344
DB 224 HPASSTKVYDKKIEPRGPTIKP--CPPCKCPAPNLLGGPSVFIFFPKIKDVLMISSPMT 281
QY 345 CVVVDVSHEDPEVKFNWYVDGVGVVHNAKTKPREQYNSTYRVSVLTVLHODMLNGKEYK 404
DB 282 CVVVDVSEDDPDVQISWPNVNVVLTAQTQTHREDYNSTLRFVSALPIQHDWMSGKEFK 341
QY 405 CKVSNKALPVPIEKTISKAKGQPREPQVYTLPPSREEETKQVSLTCLVKGFPSPD'AVE 464
DB 342 CKVNNKALPAPIERTISKPKGSVRAPQVYVLPPEEEMTKKQVTLTCKVTDFFMPEDIYVE 401
QY 465 WESNGQPNNTKTTTPVLDSGSPFLYSKLTVDKSRQOQGNVFSQVWHEALHNHYTKS 524
DB 402 WTNNGKTELNYKNTEPVLDSGSPFLYSKLTVDKSRQOQGNVFSQVWHEALHNHYTKS 461
QY 525 LSLSPGK 531
DB 462 FSRTPGK 468
```

Search completed: October 14, 2003, 06:19:32

Job time : 58.8167 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:16:09 ; Search time 314.154 Seconds  
(without alignments)  
268.288 Million cell updates/sec

Title: US-10-068-426-5

Perfect score: 2839

Sequence: 1 MPILLLLLLPSPLRPHPIC.....MHEALHNHYTKQSLSPGK 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2839	100.0	531	23	ABB78238 Amino acid sequenc
2	2829	99.6	531	23	ABB78237 Amino acid sequenc
3	2822.5	99.4	544	23	ABB78236 Amino acid sequenc
4	2822	99.4	531	23	ABB78239 Amino acid sequenc
5	2812.5	99.1	544	23	ABB78234 Amino acid sequenc
6	2812.5	99.1	544	23	ABB78235 Amino acid sequenc
7	2411.5	84.9	562	21	AAV49933 Human glycoprotein
8	2381.5	83.9	568	21	AAV49935 Human glycoprotein
9	1615.5	56.9	626	22	AAE12155 Human glycoprotein

10	1536	54.1	250	23	ABB78244	Amino acid sequenc
11	1527	53.8	301	23	ABB78240	Amino acid sequenc
12	1526	53.8	290	23	ABB78243	Amino acid sequenc
13	1526	53.8	293	10	AA291368	45 kDa amino termi
14	1521.5	53.6	302	23	ABB78241	Amino acid sequenc
15	1521.5	53.6	610	17	AA89436	Mutated platelet g
16	1521.5	53.6	610	18	AAW18201	Platelet glycoprot
17	1520	53.5	290	23	ABB78245	Amino acid sequenc
18	1518.5	53.5	610	15	AA51116	Platelet glycoprot
19	1514.5	53.3	610	15	AA56664	Mutant platelet g
20	1513.5	53.3	300	23	ABB78242	Amino acid sequenc
21	1237	43.6	313	19	AA53321	P-selectin ligand
22	1237	43.6	313	20	AA29766	P-selectin ligand
23	1237	43.6	313	23	AA29766	P-selectin ligand
24	1235.5	43.5	420	23	AA29766	Human dimeric P-se
25	1234	43.5	778	18	AAW15489	Human erythropoiet
26	1230.5	43.3	401	18	AAW10537	Pemphigus foliaceu
27	1230	43.3	388	22	AAW50248	Leptin 1-167/IGG1
28	1230	43.3	449	20	AA86278	Human interleukin-
29	1230	43.3	502	24	ABJ37109	Human noggin/immun
30	1229.5	43.3	542	23	AAU76357	Concatermic immun
31	1228.5	43.3	528	23	AAU7654	Fc disintegrin con
32	1226.5	43.2	397	18	AAW24060	Human soluble PSGl
33	1226.5	43.2	397	18	AAW22722	Human obesity prot
34	1226.5	43.2	399	18	AAW10536	Human obesity prot
35	1225.5	43.2	480	20	AAW90206	Leptin 1-167/IGG1
36	1225.5	43.2	480	24	ABU07263	hB7.1Fc soluble fu
37	1225.5	43.2	541	23	AAE29077	Human expressed pr
38	1225.5	43.2	558	23	AAE29076	Human IL-22R-mutat
39	1225.5	43.2	580	23	AAE23302	Human nectin-4-IGG
40	1225	43.1	476	22	AA49243	Chimeric 4H6 anti-
41	1225	43.1	451	20	AAV50031	Human E27 anti-IgE
42	1223	43.1	451	20	AAW95859	Mus musculus anti-
43	1223	43.1	451	20	AAW95861	Mus musculus anti-
44	1223	43.1	451	20	AAW95663	Mus musculus anti-
45	1223	43.1	451	21	AA807473	Amino acid sequenc

#### ALIGNMENTS

#### RESULT 1

ABB78238

ID ABB78238 standard; Protein; 531 AA.

AC ABB78238;

DT 25-NOV-2002 (first entry)

DE Amino acid sequence of GPIIb290/2V-Ig fusion protein.

KW Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;

KW leukocyte; platelet activation; ischaemic heart disease; atherosclerosis;

KW acute myocardial infarction; stroke; venous thrombosis; thrombocytopenia;

KW arterial thrombosis; angina; vascular condition; inflammation;

KW thrombosis; angioplasty; restenosis.

OS Synthetic.

FH Key Location/Qualifiers

FT Misc-difference 487 /note= "Ser encoded by CCC"

PN WO200263003-A2.

PD 15-AUG-2002.

PF 06-FEB-2002; 2002WO-US03549.

PR 06-FEB-2001; 2001US-266838P.

PA (GEM ) GENETICS INST LLC.

XX



PI Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
 XX WPI; 2002-657537/70.  
 DR N-PSDB; ABQ78665.  
 XX  
 PT New glycoprotein Ib alpha fusion polypeptides, useful for treating a  
 PT disorder associated with platelet activation e.g. ischaemic heart  
 PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
 XX  
 PS Claim 20; Page 3-4; 45pp; English.  
 XX  
 CC The present sequence represents a fusion protein of glycoprotein  
 CC IB-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion  
 CC protein inhibits the adherence of platelets to leukocytes. The fusion  
 CC polypeptide is useful for treating a disorder associated with platelet  
 CC activation e.g. ischaemic heart disease, acute myocardial infarction,  
 CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
 CC unstable angina. It can also be used to treat vascular conditions  
 CC associated with vascular inflammation, thrombosis, and  
 CC angioplasty-related restenosis.  
 XX  
 SQ Sequence 531 AA;  
 Query Match 100.0%; Score 2839; DB 23; Length 531;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-204;  
 Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MPLLILLLLLPSPLPHPICEVSVASHLEVNCNRLTALPPDLPKDTILHSENLLY 60  
 Db 1 MPLLILLLLLPSPLPHPICEVSVASHLEVNCNRLTALPPDLPKDTILHSENLLY 60  
 QY 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPLVGLTLDLSHNQQLSLPLGOTLPALT 120  
 Db 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPLVGLTLDLSHNQQLSLPLGOTLPALT 120  
 QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTTELP 180  
 Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTTELP 180  
 QY 181 AGLNGLENLDTLLQENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240  
 Db 181 AGLNGLENLDTLLQENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240  
 QY 241 ENVYWKQVVDKAVTSNVASVQCDNSDKFPVYKPKGKPTLGDGDTLDLYYPEEDT 300  
 Db 241 ENVYWKQVVDKAVTSNVASVQCDNSDKFPVYKPKGKPTLGDGDTLDLYYPEEDT 300  
 QY 361 WYVDGVEVHNAKTPREQYNSTYRVVSVLTVLHQDLNKGKCKVSNKALPVPIEKT 420  
 Db 361 WYVDGVEVHNAKTPREQYNSTYRVVSVLTVLHQDLNKGKCKVSNKALPVPIEKT 420  
 QY 421 SKAKGQPREQVYTLPPSREMTKNQVSLTCLVKGFPDSQIAVWESNGQFPENNYKTPP 480  
 Db 421 SKAKGQPREQVYTLPPSREMTKNQVSLTCLVKGFPDSQIAVWESNGQFPENNYKTPP 480  
 QY 481 VLDSDGSFFLYSKLTVDSKRWQGNVFSVCSVMHEALHNNHYTKLSLSLSPGK 531  
 Db 481 VLDSDGSFFLYSKLTVDSKRWQGNVFSVCSVMHEALHNNHYTKLSLSLSPGK 531

RESULT 2  
 ABB78237  
 ID ABB78237 standard; Protein; 531 AA.  
 XX  
 AC ABB78237;  
 XX  
 DT 25-NOV-2002 (first entry)  
 XX  
 DE Amino acid sequence of GPIb290-Ig fusion protein.

XX Glycoprotein IB-alpha; GPIb; immunoglobulin; Ig; platelet adherence;  
 KW leukocyte; platelet activation; ischaemic heart disease;  
 KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
 KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
 XX thrombosis; angioplasty; restenosis.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 487 /note= "Ser encoded by CCC"  
 XX  
 PN WO200263003-A2.  
 XX 15-AUG-2002.  
 PD  
 XX 06-FEB-2002; 2002WO-US03549.  
 PF  
 XX 06-FEB-2001; 2001US-266838P.  
 PR  
 XX (GEMY ) GENETICS INST LLC.  
 PA  
 XX Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
 PI WPI; 2002-657537/70.  
 XX N-PSDB; ABQ78664.  
 DR  
 XX New glycoprotein Ib alpha fusion polypeptides, useful for treating a  
 PT disorder associated with platelet activation e.g. ischaemic heart  
 PT disease, stroke, venous or arterial thrombosis or atherosclerosis -  
 XX  
 PS Claim 20; Page 3; 45pp; English.  
 XX  
 CC The present sequence represents a fusion protein of glycoprotein  
 CC IB-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion  
 CC protein inhibits the adherence of platelets to leukocytes. The fusion  
 CC polypeptide is useful for treating a disorder associated with platelet  
 CC activation e.g. ischaemic heart disease, acute myocardial infarction,  
 CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
 CC unstable angina. It can also be used to treat vascular conditions  
 CC associated with vascular inflammation, thrombosis, and  
 CC angioplasty-related restenosis.  
 XX  
 SQ Sequence 531 AA;  
 Query Match 99.6%; Score 2829; DB 23; Length 531;  
 Best Local Similarity 99.6%; Pred. No. 4.4e-203;  
 Matches 529; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MPLLILLLLLPSPLPHPICEVSVASHLEVNCNRLTALPPDLPKDTILHSENLLY 60  
 Db 1 MPLLILLLLLPSPLPHPICEVSVASHLEVNCNRLTALPPDLPKDTILHSENLLY 60  
 QY 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPLVGLTLDLSHNQQLSLPLGOTLPALT 120  
 Db 61 TFSLATLMPYTRLTQNLNDRCELTKLQVDGTLPLVGLTLDLSHNQQLSLPLGOTLPALT 120  
 QY 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTTELP 180  
 Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTTELP 180  
 QY 181 AGLNGLENLDTLLQENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240  
 Db 181 AGLNGLENLDTLLQENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEILYFRWLQDNA 240  
 QY 241 ENVYWKQVVDKAVTSNVASVQCDNSDKFPVYKPKGKPTLGDGDTLDLYYPEEDT 300  
 Db 241 ENVYWKQVVDKAVTSNVASVQCDNSDKFPVYKPKGKPTLGDGDTLDLYYPEEDT 300  
 QY 301 EGDVVRPHTCPCPAPENALGAPSVFLPFPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 360  
 Db 301 EGDVVRPHTCPCPAPENALGAPSVFLPFPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 360

QY 361 WYVDGVEVHNATKPREEOYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPVPPIEKTI 420  
 DB 361 WYVDGVEVHNATKPREEOYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPVPPIEKTI 420  
 QY 421 SKAKGPREPQVYTLPPSREEMTKNQVSLTCLVKGFPYSDIAVWESNGOPENNYKTTTP 480  
 DB 421 SKAKGPREPQVYTLPPSREEMTKNQVSLTCLVKGFPYSDIAVWESNGOPENNYKTTTP 480  
 QY 481 VLDSGDSFFLYSKLTVDKSRWQGNVFCVSVWHEALHNHYTKCSLSLSPGK 531  
 DB 481 VLDSGDSFFLYSKLTVDKSRWQGNVFCVSVWHEALHNHYTKCSLSLSPGK 531

RESULT 3  
 ABB78236  
 ID ABB78236 standard; Protein; 544 AA.  
 XX  
 AC ABB78236;  
 XX  
 DT 25-NOV-2002 (first entry)  
 XX  
 DE Amino acid sequence of GP1b302/4X-Ig fusion protein.  
 XX  
 KW Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;  
 KW leukocyte; platelet activation; ischaemic heart disease;  
 KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
 KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
 KW thrombosis; angioplasty; restenosis.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 500 /note= "Ser encoded by CCC"  
 FT  
 XX  
 PN WO200263003-A2.  
 XX  
 PD 15-AUG-2002.  
 XX  
 PF 06-FEB-2002; 2002WO-US03549.  
 XX  
 PR 06-FEB-2001; 2001US-266838P.  
 XX  
 PA (GEMY ) GENETICS INST LLC.  
 XX  
 PI Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;  
 XX  
 DR WPI; 2002-657537/70.  
 DR N-PSDB; ABQ78663.  
 XX  
 PT New glycoprotein 1b alpha fusion polypeptides, useful for treating a  
 disorder associated with platelet activation e.g. ischaemic heart  
 disease, stroke, venous or arterial thrombosis or atherosclerosis -  
 XX  
 PS Claim 20; Page 3; 45pp; English.  
 XX  
 CC The present sequence represents a fusion protein of glycoprotein  
 1B-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion  
 protein inhibits the adherence of platelets to leukocytes. The fusion  
 polypeptide is useful for treating a disorder associated with platelet  
 activation e.g. ischaemic heart disease, acute myocardial infarction,  
 stroke, venous thrombosis, atherosclerosis, arterial thrombosis or  
 unstable angina. It can also be used to treat vascular conditions  
 associated with vascular inflammation, thrombosis, and  
 angioplasty-related restenosis.  
 CC  
 XX  
 SQ Sequence 544 AA;

Query Match 99.4%; Score 2822.5; DB 23; Length 544;  
 Best Local Similarity 97.6%; Pred. No. 1.4e-202;  
 Matches 531; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

QY 1 MPLLLLLLLPSPLPHIPICEVSKVASHLEVNCNRLTALPPDLPKDPTTLHLSNLLY 60  
 DB 1 MPLLLLLLLPSPLPHIPICEVSKVASHLEVNCNRLTALPPDLPKDPTTLHLSNLLY 60  
 QY 61 TFSLATLMPYTRITQNLDRCELTKLQVDTGLPVLTGLDLSHNQLQSLPLLGOTLPALT 120  
 DB 61 TFSLATLMPYTRITQNLDRCELTKLQVDTGLPVLTGLDLSHNQLQSLPLLGOTLPALT 120  
 QY 121 LDVSENLRLSLPLGALRGELGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180  
 DB 121 LDVSENLRLSLPLGALRGELGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLP 180  
 QY 181 AGLLNGLENLDTLLQENSILYTIIPKGFSGSHLLPFAFLHGNPMLCNCEILYFRWLQDNA 240  
 DB 181 AGLLNGLENLDTLLQENSILYTIIPKGFSGSHLLPFAFLHGNPMLCNCEILYFRWLQDNA 240  
 QY 241 ENYVWKQVVDVKAVTSNVASVQCDNSDKPPVYKYKPGKGCPTLGDGEGDTLDLYYPEEDT 300  
 DB 241 ENYVWKQVVDVKAVTSNVASVQCDNSDKPPVYKYKPGKGCPTLGDGEGDTLDLYYPEEDT 300  
 QY 301 EGDKV-----RPHTCPCPAPEALGAPSVFLFPKPKDPTLMISRTPEVTCV 347  
 DB 301 EGDKV-----RPHTCPCPAPEALGAPSVFLFPKPKDPTLMISRTPEVTCV 347  
 QY 348 VDVSHPEDPEVKFNWYVDGVEVHNATKPREEQNSTYRVSVLTVLHODWLNKGEYCKV 407  
 DB 348 VDVSHPEDPEVKFNWYVDGVEVHNATKPREEQNSTYRVSVLTVLHODWLNKGEYCKV 407  
 QY 408 SNKALPVPPIEKTISKAGOPREPQVYTLPPSREEMTKNQVSLTCLVKGFPYSDIAVWES 467  
 DB 408 SNKALPVPPIEKTISKAGOPREPQVYTLPPSREEMTKNQVSLTCLVKGFPYSDIAVWES 467  
 QY 468 NGOPENNYKTTTPVLDSGDSFFLYSKLTVDKSRWQGNVFCVSVWHEALHNHYTKCSLSL 527  
 DB 468 NGOPENNYKTTTPVLDSGDSFFLYSKLTVDKSRWQGNVFCVSVWHEALHNHYTKCSLSL 527  
 QY 528 SPGK 531  
 DB 541 SPGK 544

RESULT 4  
 ABB78239  
 ID ABB78239 standard; Protein; 531 AA.  
 XX  
 AC ABB78239;  
 XX  
 DT 25-NOV-2002 (first entry)  
 XX  
 DE Amino acid sequence of GP1b290/1A-Ig fusion protein.  
 XX  
 KW Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;  
 KW leukocyte; platelet activation; ischaemic heart disease;  
 KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;  
 KW arterial thrombosis; angina; vascular condition; vascular inflammation;  
 KW thrombosis; angioplasty; restenosis.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 253 /note= "Val encoded by GCG"  
 FT Misc-difference 487 /note= "Ser encoded by CCC"  
 FT  
 XX  
 PN WO200263003-A2.  
 XX  
 PD 15-AUG-2002.  
 XX  
 PF 06-FEB-2002; 2002WO-US03549.  
 XX  
 PR 06-FEB-2001; 2001US-266838P.  
 XX

```

2A (GEMY ) GENETICS INST LLC.
XX Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;
XX WPI: 2002-657537/70.
XX N-PSDB; ABQ78666.
XX New glycoprotein Ib alpha fusion polypeptides, useful for treating a
XX disorder associated with platelet activation e.g. ischaemic heart
XX disease, stroke, venous or arterial thrombosis or atherosclerosis -
XX Claim 20; Page 4; 45pp; English.
XX
XX The present sequence represents a fusion protein of glycoprotein
XX IB-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion
XX protein inhibits the adherence of platelets to leukocytes. The fusion
XX polypeptide is useful for treating a disorder associated with platelet
XX activation e.g. ischaemic heart disease, acute myocardial infarction,
XX stroke, venous thrombosis, atherosclerosis, arterial thrombosis or
XX unstable angina. It can also be used to treat vascular conditions
XX associated with vascular inflammation, thrombosis, and
XX angioplasty-related restenosis.
XX
XX Sequence 531 AA;
XX
XX Query Match 99.4%; Score 2822; DB 23; Length 531;
XX Best Local Similarity 99.4%; Pred. No. 1.5e-202;
XX Matches 528; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 MLLLLLLLLLSPHPHPICVSVKASHLEVNCNDRNLTPDLPKDTTLHLSNLLY 60
DB 1 MLLLLLLLLLSPHPHPICVSVKASHLEVNCNDRNLTPDLPKDTTLHLSNLLY 60
QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQLSLPLGOTLPALT 120
DB 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQLSLPLGOTLPALT 120
QY 121 LDVSNRNLTSPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
DB 121 LDVSNRNLTSPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
QY 181 AGLNGLENLDTLLQENSLYTIKPGFGSHLLPFAFHGNPMLCNCEIIFYFRWLQDNA 240
DB 181 AGLNGLENLDTLLQENSLYTIKPGFGSHLLPFAFHGNPMLCNCEIIFYFRWLQDNA 240
QY 241 ENVYVWKQGVVDKAVTSNVASVQCDNSDKFPVYKPGKCPDGLDEGDTLDLYYPEEDT 300
DB 241 ENVYVWKQGVVDKAVTSNVASVQCDNSDKFPVYKPGKCPDGLDEGDTLDLYYPEEDT 300
QY 301 EGDKVRPHTCPPCPAPEALGAPSVFLFPPKPKDTLMISRPTVTCVVVDVSHEDPEVKFN 360
DB 301 EGDKVRPHTCPPCPAPEALGAPSVFLFPPKPKDTLMISRPTVTCVVVDVSHEDPEVKFN 360
QY 361 WYVDGVEVHNKTKPREEQYNSYTRVSVLTVLQDHALNGYKCKVSNKALPVP-EKTI 420
DB 361 WYVDGVEVHNKTKPREEQYNSYTRVSVLTVLQDHALNGYKCKVSNKALPVP-EKTI 420
QY 421 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 480
DB 421 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 480
QY 481 VLDSGGSFLLYSKLTVDKSRWQQGNVFNCSVMHEALHNHYTQKSLSISPGK 531
DB 481 VLDSGGSFLLYSKLTVDKSRWQQGNVFNCSVMHEALHNHYTQKSLSISPGK 531

```

RESULT 5

ABB78234

ID ABB78234 standard; Protein; 544 AA.

XX ABB78234;

XX ABB78234;

DT 25-NOV-2002 (first entry)

```

XX Amino acid sequence of GPIb302-Ig fusion protein.
XX Glycoprotein IB-alpha; GPIb; immunoglobulin; Ig; platelet adherence;
XX leukocyte; platelet activation; ischaemic heart disease;
XX acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;
XX arterial thrombosis; angina; vascular condition; vascular inflammation;
XX thrombosis; angioplasty; restenosis.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 500 /note= "Ser encoded by CCC"
XX WO200263003-A2.
XX 15-AUG-2002.
XX 06-FEB-2002; 2002WO-US03549.
XX 06-FEB-2001; 2001US-266838P.
XX (GEMY ) GENETICS INST LLC.
XX Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;
XX WPI: 2002-657537/70.
XX N-PSDB; ABQ78661.
XX
XX New glycoprotein Ib alpha fusion polypeptides, useful for treating a
XX disorder associated with platelet activation e.g. ischaemic heart
XX disease, stroke, venous or arterial thrombosis or atherosclerosis -
XX Claim 5; Page 2; 45pp; English.
XX
XX The present sequence represents a fusion protein of glycoprotein
XX IB-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion
XX protein inhibits the adherence of platelets to leukocytes. The fusion
XX polypeptide is useful for treating a disorder associated with platelet
XX activation e.g. ischaemic heart disease, acute myocardial infarction,
XX stroke, venous thrombosis, atherosclerosis, arterial thrombosis or
XX unstable angina. It can also be used to treat vascular conditions
XX associated with vascular inflammation, thrombosis, and
XX angioplasty-related restenosis.
XX
XX Sequence 544 AA;
XX
XX Query Match 99.1%; Score 2812.5; DB 23; Length 544;
XX Best Local Similarity 97.2%; Pred. No. 7.9e-202;
XX Matches 529; Conservative 1; Mismatches 1; Indels 13; Gaps 1;
XX
QY 1 MLLLLLLLLLSPHPHPICVSVKASHLEVNCNDRNLTPDLPKDTTLHLSNLLY 60
DB 1 MLLLLLLLLLSPHPHPICVSVKASHLEVNCNDRNLTPDLPKDTTLHLSNLLY 60
QY 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQLSLPLGOTLPALT 120
DB 61 TFSLATLMPYTRLTQNLDRCELTKLQVDTGLPVLTGLDLSHNQLSLPLGOTLPALT 120
QY 121 LDVSNRNLTSPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
DB 121 LDVSNRNLTSPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLP 180
QY 181 AGLNGLENLDTLLQENSLYTIKPGFGSHLLPFAFHGNPMLCNCEIIFYFRWLQDNA 240
DB 181 AGLNGLENLDTLLQENSLYTIKPGFGSHLLPFAFHGNPMLCNCEIIFYFRWLQDNA 240
QY 241 ENVYVWKQGVVDKAVTSNVASVQCDNSDKFPVYKPGKCPDGLDEGDTLDLYYPEEDT 300
DB 241 ENVYVWKQGVVDKAVTSNVASVQCDNSDKFPVYKPGKCPDGLDEGDTLDLYYPEEDT 300
QY 301 EGDKV-----RPHTCPPCPAPEALGAPSVFLFPPKPKDTLMISRPTVTCVV 347

```

```

|||||
301 EGDYRATRTVVKPPTKARPHTCPPCPAPEALGAPSVLFPPPKDPTLMISRTPEVTCVV 360
348 VDVSHPEDPEVKFNMYDGVGVHNAKTKPREEQYNSTYRVVSVLTVLHQDMLNGKEYKCKV 407
361 VDVSHPEDPEVKFNMYDGVGVHNAKTKPREEQYNSTYRVVSVLTVLHQDMLNGKEYKCKV 420
408 SNKALPVPPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEMES 467
421 SNKALPVPPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEMES 480
468 NGQPENNYKTTTPPVLDSGDSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL 527
481 NGQPENNYKTTTPPVLDSGDSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL 540
528 SPKG 531
541 SPKG 544

RESULT 6
ABB78235
ID ABB78235 standard; Protein; 544 AA.
XX AC ABB78235;
XX DT 25-NOV-2002 (first entry)
XX DE Amino acid sequence of GPIb302/2A-Ig fusion protein.
XX KW Glycoprotein Ib-alpha; GPIb; immunoglobulin; Ig; platelet adherence;
XX KW leukocyte; platelet activation; ischaemic heart disease;
XX KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;
XX KW arterial thrombosis; angina; vascular condition; vascular inflammation;
XX KW thrombosis; angioplasty; restenosis.
XX OS Synthetic.
XX PH Key Location/Qualifiers
XX FT Misc-difference 500
XX FT /note= "Ser encoded by CCC"
XX PN WO200263003-A2.
XX PD 15-AUG-2002.
XX PF 06-FEB-2002; 2002WO-US03549.
XX PR 06-FEB-2001; 2001US-266838P.
XX PA (GEMY ) GENETICS INST LLC.
XX PI Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;
XX DR WPI; 2002-657537/70.
XX DR N-PSDB; ABQ78662.
XX PT New glycoprotein Ib alpha fusion polypeptides, useful for treating a
XX PT disorder associated with platelet activation e.g. ischaemic heart
XX PT disease, stroke, venous or arterial thrombosis or atherosclerosis -
XX PS Claim 20; Page 3; 45pp; English.
XX CC The present sequence represents a fusion protein of glycoprotein
XX CC 1B-alpha (GPIb) and an immunoglobulin (Ig) polypeptide. The fusion
XX CC protein inhibits the adherence of platelets to leukocytes. The fusion
XX CC polypeptide is useful for treating a disorder associated with platelet
XX CC activation e.g. ischaemic heart disease, acute myocardial infarction,
XX CC stroke, venous thrombosis, atherosclerosis, arterial thrombosis or
XX CC unstable angina. It can also be used to treat vascular conditions
XX CC associated with vascular inflammation, thrombosis, and
XX CC angioplasty-related restenosis.

```

```

SQ Sequence 544 AA;
Query Match 99.1%; Score 2812.5; DB 23; Length 544;
Best Local Similarity 97.2%; Pred. No. 7.9e-202;
Matches 529; Conservative 1; Mismatches 1; Indels 13; Gaps 1;

QY 1 MPTLLLLLLPSPLPHPHPICEVSKVASHLEVNCNRLTALPPDLPKDTTILHSENLLY 60
DB 1 MPTLLLLLLPSPLPHPHPICEVSKVASHLEVNCNRLTALPPDLPKDTTILHSENLLY 60
QY 61 TFSLATLMPYTRLTQLNDRCELTKLQVGTPLVLTGLDLSHNLQSLPLIGOTLPALTV 120
DB 61 TFSLATLMPYTRLTQLNDRCELTKLQVGTPLVLTGLDLSHNLQSLPLIGOTLPALTV 120
QY 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLKLKSLANNLTPL 180
DB 121 LDVSFNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLKLKSLANNLTPL 180
QY 181 AGLLNGLENLDTLLLOENSLYTIIPKFGFSGHLLPFAFLHGNPWLNCCEILYFRRLQDNA 240
DB 181 AGLLNGLENLDTLLLOENSLYTIIPKFGFSGHLLPFAFLHGNPWLNCCEILYFRRLQDNA 240
QY 241 ENVYVWKQVVDVKAVTSNVAQVQCDNSDKFPVYKYPKGCTPLGDEGDTLDLYDYPEEDT 300
DB 241 ENVYVWKQVVDVKAVTSNVAQVQCDNSDKFPVYKYPKGCTPLGDEGDTLDLYDYPEEDT 300
QY 301 EGDKV-----RPHTCPPCPAPEALGAPSVLFPPPKDPTLMISRTPEVTCVV 347
DB 301 EGDKVAATATVVKPPTKARPHTCPPCPAPEALGAPSVLFPPPKDPTLMISRTPEVTCVV 360
QY 348 VDVSHPEDPEVKFNMYDGVGVHNAKTKPREEQYNSTYRVVSVLTVLHQDMLNGKEYKCKV 407
DB 361 VDVSHPEDPEVKFNMYDGVGVHNAKTKPREEQYNSTYRVVSVLTVLHQDMLNGKEYKCKV 420
QY 408 SNKALPVPPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEMES 467
DB 421 SNKALPVPPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEMES 480
QY 468 NGQPENNYKTTTPPVLDSGDSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL 527
DB 481 NGQPENNYKTTTPPVLDSGDSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL 540
QY 528 SPKG 531
DB 541 SPKG 544

RESULT 7
AAV49933
ID AAV49933 standard; Protein; 562 AA.
XX AC AAV49933;
XX DT 01-FEB-2000 (first entry)
XX DE Human glycoprotein Ib/mouse IgG1Fc chimeric protein.
XX KW Glycoprotein Ib; glycoalkalidin; detection; antithrombotic; binding;
XX KW von Willebrand factor; botrocetin; chimeric protein; immunoglobulin;
XX KW thrombotic disease; diagnosis; cardiac infarction; cerebral embolism.
XX OS Chimeric - Homo sapiens.
XX OS Chimeric - Mus musculus.
XX FH Key Location/Qualifiers
XX FT Peptide 1..16 /label= signal
XX FT Protein 17..562 /label= human glycoprotein_Ib/mouse_IgG1Fc_chimeric
XX FT protein
XX PN WO9954360-A1.
XX

```

```

PD 28-OCT-1999.
PF 13-JAN-1999; 99WO-JP00089.
PR 23-APR-1998; 98JP-0113962.
PA (AJIN ) AJINOMOTO CO INC.
PI Fukuchi N, Futaki F, Kito M, Sato S, Kajiuura T, Ono Y, Ishii K;
PI Tanaka A, Shinozaki J, Jojima Y;
XX WPI; 2000-013233/01.
DR N-PSDB; AAZ35701.
XX
XX Direct quantitative detection of glycoalkaloid with immobilized von
PT Willebrand factor to bond with chimeric protein via inhibiting
PT glycoprotein Ib binding, for diagnosis of thrombotic diseases and
PT screening anti-thrombotic substances -
XX
XX Example 1; Page 70-71; 83pp; Japanese.
XX
XX A method has been developed for the detection of binding between the von
CC Willebrand factor and glycoprotein Ib or of the binding inhibition, in
CC which the von Willebrand factor immobilised in a reactor reacts with
CC glycoprotein Ib in the presence of a binding inducer to promote binding
CC between the von Willebrand factor and glycoprotein Ib. This method is
CC for the detection of glycoalkaloid as a means of thrombotic disease
CC diagnosis e.g. for cardiac infarction and cerebral embolism, and also
CC for screening substances with anti-thrombotic activity for the
CC prevention and treatment of thrombotic diseases. The method is direct,
CC convenient and quantitative, with reproducibility, and there is no need
CC to construct a monoclonal antibody for the assay. The present sequence
CC is a human glycoprotein Ib/mouse immunoglobulin gamma 1 Fc chimeric
CC protein from the present invention.
XX
SQ Sequence 562 AA;
Query Match 84.9%; Score 2411.5; DB 21; Length 562;
Best Local Similarity 79.5%; Pred. No. 8e-172;
Matches 449; Conservative 40; Mismatches 39; Indels 37; Gaps 4;
QY 1 MPELLLLLLLSPHPHPCVSVASHLEVNCNRLTALPPDLPKDTILHLSENLLY 60
DB 1 MPELLLLLLLSPHPHPCVSVASHLEVNCNRLTALPPDLPKDTILHLSENLLY 60
QY 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLPLVGLTLDLSHNOQLPLGGTLPALT 120
DB 61 TFSLATLMPYTRLTQNLDRCELTKLOVDGTLPLVGLTLDLSHNOQLPLGGTLPALT 120
QY 121 LDVSNRLTSLPLGALRGELGLOELYLKGNELKTLPPGLTTPPKLEKLSLANNLTPEL 180
DB 121 LDVSNRLTSLPLGALRGELGLOELYLKGNELKTLPPGLTTPPKLEKLSLANNLTPEL 180
QY 181 AGLINGLENLTLLOENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEILYFRRLQDNA 240
DB 181 AGLINGLENLTLLOENSLYTIKPGFGSHLLPFAFLHGNPWLNCNCEILYFRRLQDNA 240
QY 241 ENVYVWKQVDVAVTSNVASVQCDNSDKFPVKYKPGKGCPTLDGEGDTLDYIPEEDT 300
DB 241 ENVYVWKQVDVAVTSNVASVQCDNSDKFPVKYKPGKGCPTLDGEGDTLDYIPEEDT 300
QY 301 EGDKVR-----PHT-----CPP--CPAPEALGAPSVFL 326
DB 301 EGDKVRATRTVVKPPTKARTTPWGLFYSWSTASLDVPRDCGCKPCICTVPE---VSSVFI 357
QY 327 FPPKPKDTLMSRTPEVTCVVVDVSHEDPEVKFVWYDGVVEVHNAKTKPREEQVNSTYRV 386
DB 327 FPPKPKDTLMSRTPEVTCVVVDVSHEDPEVKFVWYDGVVEVHNAKTKPREEQVNSTYRV 386
QY 387 VSVLTVLHODLWNGKEYCKYKSNKALVPVIEKTKISKAGQPREPQVYTLPPSREEMTKNQ 446
DB 387 VSVLTVLHODLWNGKEYCKYKSNKALVPVIEKTKISKAGQPREPQVYTLPPSREEMTKNQ 446
QY 418 VSELPIMQDMLNGKEFKRCVNSAAPPAPISKTKGRPKAPQVVTIIPPKQEQMAQDK 477
DB 418 VSELPIMQDMLNGKEFKRCVNSAAPPAPISKTKGRPKAPQVVTIIPPKQEQMAQDK 477

```

```

QY 447 VSLTCLVKGFYPSDIAVWESNGOPENNYKTTTPVLDSDGSFFLYSKLTVDKSWOQGNV 506
DB 478 VSLTCLVKGFYPSDIAVWESNGOPENNYKTTTPVLDSDGSFFLYSKLTVDKSWOQGNV 537
QY 507 FSCSVWHEALHNHYTQKSLSLSPGK 531
DB 538 FTCSVLHGLEHNTKSLSPGK 562
XX
XX RESULT 8
XX ID AAY49935 standard; Protein; 568 AA.
XX AC AAY49935;
XX DT 01-FEB-2000 (first entry)
XX DE Human glycoprotein Ib/mouse IgG2aFc chimeric protein #2.
XX KW Glycoprotein Ib; glycoalkaloid; detection; antithrombotic; binding;
XX KW von Willebrand factor; botrocetin; chimeric protein; immunoglobulin;
XX KW thrombotic disease; diagnosis; cardiac infarction; cerebral embolism.
XX OS Chimeric - Homo sapiens.
XX OS Chimeric - Mus musculus.
XX PH Key Location/Qualifiers
XX FT Peptide 1..16
XX FT Protein 17..568
XX FT /label= signal
XX FT /label= Human glycoprotein_Ib/mouse_IgG2aFc_chimeric
XX FT protein
XX PN WO9954360-A1.
XX
XX 28-OCT-1999.
XX
XX 13-JAN-1999; 99WO-JP00089.
XX
XX 23-APR-1998; 98JP-0113962.
XX
XX (AJIN ) AJINOMOTO CO INC.
XX
XX Fukuchi N, Futaki F, Kito M, Sato S, Kajiuura T, Ono Y, Ishii K;
XX Tanaka A, Shinozaki J, Jojima Y;
XX
XX WPI; 2000-013233/01.
XX N-PSDB; AAZ35706.
XX
XX Direct quantitative detection of glycoalkaloid with immobilized von
PT Willebrand factor to bond with chimeric protein via inhibiting
PT glycoprotein Ib binding, for diagnosis of thrombotic diseases and
PT screening anti-thrombotic substances -
XX
XX Example 1; Page 77-79; 83pp; Japanese.
XX
XX A method has been developed for the detection of binding between the von
CC Willebrand factor and glycoprotein Ib or of the binding inhibition, in
CC which the von Willebrand factor immobilised in a reactor reacts with
CC glycoprotein Ib in the presence of a binding inducer to promote binding
CC between the von Willebrand factor and glycoprotein Ib. This method is
CC for the detection of glycoalkaloid as a means of thrombotic disease
CC diagnosis e.g. for cardiac infarction and cerebral embolism, and also
CC for screening substances with anti-thrombotic activity for the
CC prevention and treatment of thrombotic diseases. The method is direct,
CC convenient and quantitative, with reproducibility, and there is no need
CC to construct a monoclonal antibody for the assay. The present sequence
CC represents a human glycoprotein Ib/mouse immunoglobulin gamma 2a Fc
CC chimeric protein from the present invention.
XX
XX Sequence 568 AA;

```

Query Match 83.9%; Score 2381.5; DB 21; Length 568;



```

ABB78244
ID ABB78244 standard; Protein; 290 AA.
XX
AC ABB78244;
XX
DT 25-NOV-2002 (first entry)
XX
DE Amino acid sequence of glycoprotein 1B-alpha polypeptide GPIb290/2V.
XX
KW Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;
KW leukocyte; platelet activation; ischaemic heart disease;
KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;
KW arterial thrombosis; angina; vascular condition; vascular inflammation;
KW thrombosis; angioplasty; restenosis.
XX
OS Unidentified.
XX
PN WO200263003-A2.
XX
PD 15-AUG-2002.
XX
PF 06-FEB-2002; 2002WO-US03549.
XX
PR 06-FEB-2001; 2001US-266838P.
XX
PA (GEMY ) GENETICS INST LLC.
XX
PI Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;
XX
DR WPI; 2002-657537/70.
XX
PT New glycoprotein 1B alpha fusion polypeptides, useful for treating a
PT disorder associated with platelet activation e.g. ischaemic heart
PT disease, stroke, venous or arterial thrombosis or atherosclerosis -
XX
PS Disclosure; Page 11; 45pp; English.
XX
CC ABB78240-45 represent glycoprotein 1B-alpha (GPIb) polypeptides, which
CC are used to produce fusion proteins with an immunoglobulin (Ig)
CC polypeptide. The fusion proteins inhibit the adherence of platelets to
CC leukocytes. The fusion polypeptides are useful for treating a disorder
CC associated with platelet activation e.g. ischaemic heart disease, acute
CC myocardial infarction, stroke, venous thrombosis, atherosclerosis,
CC arterial thrombosis or unstable angina. They can also be used to treat
CC vascular conditions associated with vascular inflammation, thrombosis,
CC and angioplasty-related restenosis.
XX
SQ Sequence 290 AA;
Query Match 54.1%; Score 1536; DB 23; Length 290;
Best Local Similarity 100.0%; Pred. No. 1e-106;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHLSNLLYTFSLATMPYTRLTQ 76
Db 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHLSNLLYTFSLATMPYTRLTQ 60
QY 77 NLDRCETKLVQDGTLPVLGTLDSHNQLOSLPLGQTLPALTVLDVSFNRLTSLPLGAL 136
Db 61 NLDRCETKLVQDGTLPVLGTLDSHNQLOSLPLGQTLPALTVLDVSFNRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTETLPAGLNGLENLDTLLQ 196
Db 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTETLPAGLNGLENLDTLLQ 180
QY 197 ENSLYTTPKGFSGHLLPFAFLHGNPWLNCCEILYFRRLQDNNAENYVWKQVVDVKAVT 256
Db 181 ENSLYTTPKGFSGHLLPFAFLHGNPWLNCCEILYFRRLQDNNAENYVWKQVVDVKAVT 240
QY 257 SNVASVQCDNSDKFPVVKYPKGCGPTLGDGDTLDLYDYPPEDETEGDKVR 306
Db 241 SNVASVQCDNSDKFPVVKYPKGCGPTLGDGDTLDLYDYPPEDETEGDKVR 290

```

---

```

RESULT 11
ABB78240
ID ABB78240 standard; Protein; 301 AA.
XX
AC ABB78240;
XX
DT 25-NOV-2002 (first entry)
XX
DE Amino acid sequence of glycoprotein 1B-alpha polypeptide GPIb302.
XX
KW Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;
KW leukocyte; platelet activation; ischaemic heart disease;
KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;
KW arterial thrombosis; angina; vascular condition; vascular inflammation;
KW thrombosis; angioplasty; restenosis.
XX
OS Unidentified.
XX
PN WO200263003-A2.
XX
PD 15-AUG-2002.
XX
PF 06-FEB-2002; 2002WO-US03549.
XX
PR 06-FEB-2001; 2001US-266838P.
XX
PA (GEMY ) GENETICS INST LLC.
XX
PI Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;
XX
DR WPI; 2002-657537/70.
XX
PT New glycoprotein 1B alpha fusion polypeptides, useful for treating a
PT disorder associated with platelet activation e.g. ischaemic heart
PT disease, stroke, venous or arterial thrombosis or atherosclerosis -
XX
PS Disclosure; Page 10; 45pp; English.
XX
CC ABB78240-45 represent glycoprotein 1B-alpha (GPIb) polypeptides, which
CC are used to produce fusion proteins with an immunoglobulin (Ig)
CC polypeptide. The fusion proteins inhibit the adherence of platelets to
CC leukocytes. The fusion polypeptides are useful for treating a disorder
CC associated with platelet activation e.g. ischaemic heart disease, acute
CC myocardial infarction, stroke, venous thrombosis, atherosclerosis,
CC arterial thrombosis or unstable angina. They can also be used to treat
CC vascular conditions associated with vascular inflammation, thrombosis,
CC and angioplasty-related restenosis.
XX
SQ Sequence 301 AA;
Query Match 53.8%; Score 1527; DB 23; Length 301;
Best Local Similarity 93.0%; Pred. No. 5e-106;
Matches 292; Conservative 2; Mismatches 6; Indels 14; Gaps 1;
QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHLSNLLYTFSLATMPYTRLTQ 76
Db 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHLSNLLYTFSLATMPYTRLTQ 60
QY 77 NLDRCETKLVQDGTLPVLGTLDSHNQLOSLPLGQTLPALTVLDVSFNRLTSLPLGAL 136
Db 61 NLDRCETKLVQDGTLPVLGTLDSHNQLOSLPLGQTLPALTVLDVSFNRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTETLPAGLNGLENLDTLLQ 196
Db 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTETLPAGLNGLENLDTLLQ 180
QY 197 ENSLYTTPKGFSGHLLPFAFLHGNPWLNCCEILYFRRLQDNNAENYVWKQVVDVKAVT 256
Db 181 ENSLYTTPKGFSGHLLPFAFLHGNPWLNCCEILYFRRLQDNNAENYVWKQVVDVKAVT 240
QY 257 SNVASVQCDNSDKFPVVKYPKGCGPTLGDGDTLDLYDYPPEDETEGDKVRPHTCPPCAP 316

```



Db 241 SNVASQCDNSDKFPVYKPGKCPGLGDEGDTLDYVYPEEDTEGDKVR ----- 290

Qy 317 EALGAPSVLFPPK 330

Db 291 ----ATRTWKFPK 300

RESULT 12

ABB78243

ID ABB78243 standard; Protein; 290 AA.

XX ABB78243;

DT 25-NOV-2002 (first entry)

XX Amino acid sequence of glycoprotein 1B-alpha polypeptide GPIb290.

DE Glycoprotein 1B-alpha; GPIb; immunoglobulin; Ig; platelet adherence;

KW leukocyte; platelet activation; ischaemic heart disease;

KW acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;

KW arterial thrombosis; angina; vascular condition; vascular inflammation;

KW thrombosis; angioplasty; restenosis.

XX Unidentified.

OS

XX

XX WO200263003-A2.

XX

PD 15-AUG-2002.

XX

PF 06-FEB-2002; 2002WO-US03549.

XX

PR 06-FEB-2001; 2001US-266838P.

XX

PA (GEMY ) GENETICS INST LLC.

XX

PI Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;

XX MPI; 2002-657537/70.

XX

XX New glycoprotein 1b alpha fusion polypeptides, useful for treating a

PT disorder associated with platelet activation e.g. ischaemic heart

PT disease, stroke, venous or arterial thrombosis or atherosclerosis -

XX Disclosure; Page 10-11; 45pp; English.

PS

XX ABB78240-45 represent glycoprotein 1B-alpha (GPIb) polypeptides, which

CC are used to produce fusion proteins with an immunoglobulin (Ig)

CC polypeptide. The fusion proteins inhibit the adherence of platelets to

CC leukocytes. The fusion polypeptides are useful for treating a disorder

CC associated with platelet activation e.g. ischaemic heart disease, acute

CC myocardial infarction, stroke, venous thrombosis, atherosclerosis,

CC arterial thrombosis or unstable angina. They can also be used to treat

CC vascular conditions associated with vascular inflammation, thrombosis,

CC and angioplasty-related restenosis.

XX

XX Sequence 290 AA;

Qy Query Match 53.8%; Score 1526; DB 23; Length 290;

Best Local Similarity 99.3%; Pred. No. 5.7e-106;

Matches 288; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKOTTLHLSENLLYTFSLATLMPYTRLTQL 76

Db 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKOTTLHLSENLLYTFSLATLMPYTRLTQL 60

Qy 77 NLDRCETLKLQVGTLPVLGTLDSLNQSLPLGOTLPALTVDVSNFRLTSLPLGAL 136

Db 61 NLDRCETLKLQVGTLPVLGTLDSLNQSLPLGOTLPALTVDVSNFRLTSLPLGAL 120

Qy 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLLGNELNLTLLQ 196

Db 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLLGNELNLTLLQ 180

Qy 197 ENSLYTIPIKGFSGSHLLPFAFLHGNFWLNCCELLYFRRWLNQNAENVYVWKQVDDKAVT 256

Db 181 ENSLYTIPIKGFSGSHLLPFAFLHGNFWLNCCELLYFRRWLNQNAENVYVWKQVDDKAVT 240

Qy 257 SNVASQCDNSDKFPVYKPGKCPGLGDEGDTLDYVYPEEDTEGDKVR 306

Db 241 SNVASQCDNSDKFPVYKPGKCPGLGDEGDTLDYVYPEEDTEGDKVR 290

RESULT 13

AAP91368

ID AAP91368 standard; peptide; 293 AA.

XX AAP91368;

XX

DT 25-MAR-2003 (updated)

DT 21-MAR-1990 (first entry)

XX

XX 45 kDa amino terminal tryptic fragment of glyocalicin and derivs.

XX Glyocalicin; von Willebrand factor; platelet membrane glycoprotein 1b;

KW platelet aggregation prevention; thrombosis inhibition; antithrombotic

KW agent

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FT Peptide 18..34

FT Peptide 21..35

FT Peptide 26..40

FT Peptide 26..34

FT Peptide 141..155

FT Peptide 231..245

FT Peptide 271..285

FT Peptide 281..285

XX

XX EP317278-A.

XX

XX 24-MAY-1989.

XX

XX 16-NOV-1988; 88EP-0310799.

XX

XX 17-NOV-1987; 87US-0121454.

XX

XX (SCRI ) SCRIPPS CLINIC & RES FOUND.

XX

XX Zimmerman TS, Ruggeri ZM, Houghten RA, Vincete V, Mohri H;

XX WPI; 1989-152756/21.

XX

XX Proteolytic 45 KD fragment of glyocalicin and derivs. - which inhibit

PT binding of von Willebrand factor to platelet membrane glyco:protein, and

PT used as antithrombotic agents

XX

XX Claim 1; Page -; 10pp; English.

PS

XX The claim is for a peptide of a 45KD amino terminal tryptic fragment of

CC glyocalicin selected from the sequence which inhibits binding of von

CC Willebrand factor to platelet membrane glycoprotein 1b and related

CC molecules or other cells and cell matrices. Also claimed are a sequential

CC subset of the above (Claim 2) and specific peptides (see FT) (Claim 3)

CC with the same functions and any peptide of any sequential subset of amino

CC acids of the sequence (Claim 4). The peptides and derivs. prevent

CC platelet aggregation and inhibit thrombosis.

CC (Updated on 25-MAR-2003 to correct PA field.)

XX

XX Sequence 293 AA;

Qy Query Match 53.8%; Score 1526; DB 10; Length 293;

Best Local Similarity 99.3%; Pred. No. 5.7e-106;

Matches 288; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKOTTLHLSENLLYTFSLATLMPYTRLTQL 76

```

Db      1  HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLIYTSFSLATIMPYTRLTQ 60
      77  NLDRCCLTKLQVDTGLPVLTGLDLSHNQLOSPLLGOTLPALTVDVSNFNRLTSLPLGAL 136
      61  NLDRCCLTKLQVDTGLPVLTGLDLSHNQLOSPLLGOTLPALTVDVSNFNRLTSLPLGAL 120
      137  RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196
      121  RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180
      197  ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDNAENVYWKQVDVKAVT 256
      181  ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDNAENVYWKQVDVKAVT 240
      257  SNVASVQCDNSDKFPVYKYKPGKGCPTLGDGDTLDLYYPEEDTEGDKVR 306
      241  SNVASVQCDNSDKFPVYKYKPGKGCPTLGDGDTLDLYYPEEDTEGDKVR 290

RESULT 14
ABB78241
ID      ABB78241 standard; Protein; 302 AA.
XX      AC
XX      ABB78241;
XX
XX      25-NOV-2002 (first entry)
XX
XX      Amino acid sequence of glycoprotein Ib-alpha polypeptide GPIb302/2A.
XX      Glycoprotein Ib-alpha; GPIb; immunoglobulin; Ig; platelet adherence;
XX      leukocyte; platelet activation; ischaemic heart disease;
XX      acute myocardial infarction; stroke; venous thrombosis; atherosclerosis;
XX      arterial thrombosis; angina; vascular condition; vascular inflammation;
XX      thrombosis; angioplasty; restenosis.
XX      Unidentified.
XX
XX      WO200263003-A2.
XX
XX      15-AUG-2002.
XX
XX      06-FEB-2002; 2002WO-US03545.
XX
XX      06-FEB-2001; 2001US-266838P.
XX
XX      (GEMY ) GENETICS INST LLC.
XX
XX      Shaw GD, Kumar R, Sako DS, McDonagh T, Sullivan FX;
XX      WPI; 2002-657537/70.
XX
XX      New glycoprotein Ib alpha fusion polypeptides, useful for treating a
XX      disorder associated with platelet activation e.g. ischaemic heart
XX      disease, stroke, venous or arterial thrombosis or atherosclerosis -
XX      Disclosure; Page 10; 45pp; English.
XX
XX      ABB78240-45 represent glycoprotein Ib-alpha (GPIb) polypeptides, which
XX      are used to produce fusion proteins with an immunoglobulin (Ig)
XX      polypeptide. The fusion proteins inhibit the adherence of platelets to
XX      leukocytes. The fusion polypeptides are useful for treating a disorder
XX      associated with platelet activation e.g. ischaemic heart disease, acute
XX      myocardial infarction, stroke, venous thrombosis, atherosclerosis,
XX      arterial thrombosis or unstable angina. They can also be used to treat
XX      vascular conditions associated with vascular inflammation, thrombosis,
XX      and angioplasty-related restenosis.
XX
XX      Sequence 302 AA;
XX
XX      Query March 53.6%; Score 1521.5; DB 23; Length 302;
XX      Best Local Similarity 92.7%; Pred. No. 1.3e-105;
XX      Matches 291; Conservative 2; Mismatches 8; Indels 13; Gaps 1;

```

```

QY      17  HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLIYTSFSLATIMPYTRLTQ 76
      1  HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLIYTSFSLATIMPYTRLTQ 60
      77  NLDRCCLTKLQVDTGLPVLTGLDLSHNQLOSPLLGOTLPALTVDVSNFNRLTSLPLGAL 136
      61  NLDRCCLTKLQVDTGLPVLTGLDLSHNQLOSPLLGOTLPALTVDVSNFNRLTSLPLGAL 120
      137  RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196
      121  RGLGELQELYLKGNELKTLPPGLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180
      197  ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDNAENVYWKQVDVKAVT 256
      181  ENSLYTTPKGFPGSHLLPFAFLHGNPWLNCCEILYFRRLQDNAENVYWKQVDVKAVT 240
      257  SNVASVQCDNSDKFPVYKYKPGKGCPTLGDGDTLDLYYPEEDTEGDKVRPHTCPPCPAP 316
      241  SNVASVQCDNSDKFPVYKYKPGKGCPTLGDGDTLDLYYPEEDTEGDKV----- 289
      317  EALGAPSVLFPKP 330
      290  --AATATVVKPPTK 301

RESULT 15
AAR89436
ID      AAR89436 standard; Protein; 610 AA.
XX      AC
XX      AAR89436;
XX      25-MAR-2003 (updated)
XX      02-SEP-1996 (first entry)
XX      Mutated platelet glycoprotein-Ib-alpha GPIba protein sequence.
XX      Platelet glycoprotein-Ib-alpha; GPIba; mutagenesis; point mutation;
XX      von Willebrand factor; blood disorder; platelet disorder;
XX      protein engineering; Bernard-Soulier disease.
XX      Homo sapiens.
XX
XX      Key Location/Qualifiers
XX      Misc-difference 57
XX      /note= "substitution from Phe in wild-typ
XX      GPIb-alpha"
XX      Region 36..200
XX      /note= "Leu rich area"
XX      Region 200..220
XX      /note= "flanking region to Leu rich area"
XX      Region 220..310
XX      /note= "hinge region"
XX      Region 310..420
XX      /note= "Ser/Thr rich area"
XX
XX      US5492809-A.
XX
XX      20-FEB-1996.
XX      09-SEP-1993; 93US-0119262.
XX      15-JAN-1992; 92US-0821717.
XX      07-OCT-1991; 91US-0770968.
XX      09-SEP-1993; 93US-0119262.
XX      (UUNY ) UNIV NEW YORK STATE RES FOUND.
XX      Cunningham D, Finch CN, Lyle VA, Miller JL;
XX      WPI; 1996-128585/13.
XX      DNA encoding platelet glyco:protein Ib alpha mutant Phe57 -

```

PT introduced into platelets to reduce aggregation and reactivity with:  
PT von Willebrand factor, also probe for diagnosis of Bernard-Soulier  
PT disease

XX

PS Disclosure; Column 21-26; 20pp; English.

XX

CC A substitution of T for C at position 259 in the DNA sequence of  
CC GPIb-alpha leads to the replacement of Phe for Leu at  
CC residue 57 of the mature GPIb-alpha molecule. This mutated GPIb-  
CC alpha protein is less reactive with von Willebrand factor, a  
CC a characteristic of the autosomal recessive bleeding disorder,  
CC Bernard-Soulier disease.  
CC (Updated on 25-MAR-2003 to correct PF field.)

XX

SQ Sequence 610 AA;

```
Query Match          53.6%; Score 1521.5; DB 17; Length 610;
Best Local Similarity 82.9%; Pred. No. 3.3e-105;
Matches 301; Conservative 5; Mismatches 22; Indels 35; Gaps 5;

QY 17 HPICEVSKVASHLEVNCCKRNLTPDLPKDTTILHLSNLLYTFSLATLMPYTRLTOL 76
   |||||
Db 1 HPICEVSKVASHLEVNCCKRNLTPDLPKDTTILHLSNLLYTFSLATLMPYTRLTOL 60

QY 77 NLDRCBLTKLQVDGTLPLVGLTDLDSNQLQSLPLGQTLPALTVLDVSNRLTSLPLGAL 136
   |||||
Db 61 NLDRCBLTKLQVDGTLPLVGLTDLDSNQLQSLPLGQTLPALTVLDVSNRLTSLPLGAL 120

QY 137 RGLGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTLPAGLNGLENLDTLLLO 196
   |||||
Db 121 RGLGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTLPAGLNGLENLDTLLLO 180

QY 197 ENSLYTIPKGGFFGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNAENVYVKQVDVKAVT 256
   |||||
Db 181 ENSLYTIPKGGFFGSHLLPFAFLHGNPWLNCCEILYFRRWLQDNAENVYVKQVDVKAVT 240

QY 257 SNVASVQCDNSDKFPVYKYPGKCGPTLGDGDTLDLYDYYPEESTEGDKVR----- 306
   |||||
Db 241 SNVASVQCDNSDKFPVYKYPGKCGPTLGDGDTLDLYDYYPEESTEGDKVRATRTVVVKFPT 300

QY 307 -PHTCP-----PCPA---PEALGAPSVFLPPPK--PKDTL-----MISRTP 341
   |||
Db 301 KAHTTPWGLFYWSWTASLDQSPSSUHTQBSTKEQTTPPRWTNFTLHMESITFSKTP 360

QY 342 EVT 344
   :
Db 361 KST 363
```

Search completed: October 14, 2003, 06:30:27  
Job time : 318.154 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2003, 06:31:18 ; Search time 192.642 Seconds  
(without alignments)  
444.136 Million cell updates/sec

Title: US-10-068-426-5  
Perfect score: 2839  
Sequence: 1 MFLLLLLLLSPLPHPIC.....MHEALHNYTKSLSPGK 531

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues  
Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2839	100.0	531	15 US-10-068-426-5	Sequence 5, Appli
2	2829	99.6	531	15 US-10-068-426-4	Sequence 4, Appli
3	2823	99.4	531	15 US-10-068-426-6	Sequence 6, Appli
4	2822.5	99.4	544	15 US-10-068-426-3	Sequence 3, Appli
5	2812.5	99.1	544	15 US-10-068-426-1	Sequence 1, Appli
6	2812.5	99.1	544	15 US-10-068-426-2	Sequence 2, Appli
7	1536	54.1	230	15 US-10-068-426-11	Sequence 11, Appli
8	1526.5	53.8	302	15 US-10-068-426-7	Sequence 7, Appli
9	1526	53.6	290	15 US-10-068-426-10	Sequence 10, Appli
10	1521.5	53.6	302	15 US-10-068-426-8	Sequence 8, Appli
11	1520	53.5	290	15 US-10-068-426-12	Sequence 12, Appli
12	1514	53.3	301	15 US-10-068-426-9	Sequence 9, Appli
13	1237	43.6	313	12 US-09-825-580-4	Sequence 4, Appli
14	1237	43.6	313	15 US-10-211-786-4	Sequence 4, Appli
15	1232.5	43.4	470	15 US-10-264-634-33	Sequence 33, Appli

16	1230.5	43.3	401	9	US-09-859-361-9	Sequence 9, Appli
17	1230	43.3	388	9	US-09-784-623-16	Sequence 16, Appli
18	1230	43.3	449	16	US-10-323-268-23	Sequence 23, Appli
19	1227	43.2	467	15	US-10-264-634-29	Sequence 29, Appli
20	1226.5	43.2	397	8	US-08-779-457-47	Sequence 47, Appli
21	1226.5	43.2	399	9	US-09-859-361-7	Sequence 7, Appli
22	1225.5	43.2	382	15	US-10-207-655-307	Sequence 307, Appli
23	1225.5	43.2	580	11	US-09-972-268-36	Sequence 36, Appli
24	1223	43.1	451	9	US-09-920-171-14	Sequence 14, Appli
25	1223	43.1	451	9	US-09-920-171-16	Sequence 16, Appli
26	1223	43.1	451	9	US-09-920-171-18	Sequence 18, Appli
27	1223	43.1	451	11	US-09-925-179-65	Sequence 65, Appli
28	1223	43.1	451	11	US-09-925-179-66	Sequence 66, Appli
29	1223	43.1	451	12	US-10-113-996-14	Sequence 14, Appli
30	1223	43.1	451	12	US-10-113-996-16	Sequence 16, Appli
31	1223	43.1	451	12	US-10-113-996-18	Sequence 18, Appli
32	1223	43.1	451	12	US-10-292-869-2	Sequence 2, Appli
33	1223	43.1	451	12	US-09-792-938-2	Sequence 8, Appli
34	1222	43.0	453	9	US-09-802-077-8	Sequence 8, Appli
35	1222	43.0	453	9	US-09-802-096-8	Sequence 8, Appli
36	1222	43.0	453	11	US-09-925-179-8	Sequence 25, Appli
37	1222	43.0	467	15	US-10-264-634-25	Sequence 25, Appli
38	1222	43.0	492	15	US-10-264-634-27	Sequence 27, Appli
39	1222	43.0	533	9	US-09-792-2008-4	Sequence 4, Appli
40	1222	43.0	977	12	US-10-357-653-1	Sequence 1, Appli
41	1220.5	43.0	694	10	US-09-935-868-18	Sequence 18, Appli
42	1220.5	43.0	694	12	US-10-282-162-18	Sequence 18, Appli
43	1220.5	43.0	694	15	US-10-287-035-18	Sequence 18, Appli
44	1220.5	43.0	793	10	US-09-935-868-32	Sequence 32, Appli
45	1220.5	43.0	793	12	US-10-282-162-32	Sequence 32, Appli

ALIGNMENTS

RESULT 1  
US-10-068-426-5  
; Sequence 5, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaw, Gray D.  
; APPLICANT: Sako, Dianne S.  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Sullivan, Francis  
; APPLICANT: McDonagh, Tom  
; TITLE OF INVENTION: Platiect Glycoprotein IB Alpha Fusion Polypeptides and  
; TITLE OF INVENTION: Methods of Use Thereof  
; FILE REFERENCE: 22058-503  
; CURRENT APPLICATION NUMBER: US/10/068.426  
; CURRENT FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: 60/266,838  
; PRIOR FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(531)  
; OTHER INFORMATION: GPlb290/2V-Ig  
US-10-068-426-5

Query Match 100.0%; Score 2839; DB 15; Length 531;  
Best Local Similarity 100.0%; Pred. No. 3.9e-219;  
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MFLLLLLLLSPLPHPICVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLY 60  
DB 1 MFLLLLLLLSPLPHPICVSKVASHLEVNCCKRNLTALPPDLPKDTTILHSENLLY 60  
QY 61 TFSLATLMPYTRLTQLNLDRCLETQLQVDGTLPLVLTGLDLSHNLQSLPLGLQTLPALTV 120

```

Db      61  TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLTGLDLSHNLQSLPLLGOTLPALT 120
      |||
Qy      121  LDVSFNRUTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTSLP 180
      |||
Db      121  LDVSFNRUTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTSLP 180
      |||
Qy      181  AGLLNGLENLDTLLQENSLYTI PKGFFGSHLLPFAFLHGNPWLNCNCELILYFRWLQDNA 240
      |||
Db      181  AGLLNGLENLDTLLQENSLYTI PKGFFGSHLLPFAFLHGNPWLNCNCELILYFRWLQDNA 240
      |||
Qy      241  ENVYVMKQVVDVKAVTSNVASVQCDNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300
      |||
Db      241  ENVYVMKQVVDVKAVTSNVASVQCDNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300
      |||
Qy      301  EGDKVRPHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 360
      |||
Db      301  EGDKVRPHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 360
      |||
Qy      361  WYVDGVEVHNAKTPREQYNSYTRVVSVLTVLHQDMLNGKEYKCKVSNKALPVPLEKTI 420
      |||
Db      361  WYVDGVEVHNAKTPREQYNSYTRVVSVLTVLHQDMLNGKEYKCKVSNKALPVPLEKTI 420
      |||
Qy      421  SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 480
      |||
Db      421  SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 480
      |||
Qy      481  VLDSGSEFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTKQSLSPGK 531
      |||
Db      481  VLDSGSEFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTKQSLSPGK 531
      |||

```

## RESULT 2

```

US-10-068-426-4
; Sequence 4, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platiect Glycoprotein IB Alpha Fusion Polypeptides and
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(531)
; OTHER INFORMATION: GPIb290-IG
US-10-068-426-4

```

Query Match 99.6%; Score 2829; DB 15; Length 531;  
 Best Local Similarity 99.6%; Pred. No. 2.5e-218;  
 Matches 529; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1  MPLLLLLLLLPSPLPHPICEVSKVASHLEVNCNKNLTALPPDLPKDITILHSENLY 60
      |||
Db      1  MPLLLLLLLLPSPLPHPICEVSKVASHLEVNCNKNLTALPPDLPKDITILHSENLY 60
      |||
Qy      61  TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLTGLDLSHNLQSLPLLGOTLPALT 120
      |||
Db      61  TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLTGLDLSHNLQSLPLLGOTLPALT 120
      |||

```

```

Qy      121  LDVSFNRUTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTSLP 180
      |||
Db      121  LDVSFNRUTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTSLP 180
      |||
Qy      181  AGLLNGLENLDTLLQENSLYTI PKGFFGSHLLPFAFLHGNPWLNCNCELILYFRWLQDNA 240
      |||
Db      181  AGLLNGLENLDTLLQENSLYTI PKGFFGSHLLPFAFLHGNPWLNCNCELILYFRWLQDNA 240
      |||
Qy      241  ENVYVMKQVVDVKAVTSNVASVQCDNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300
      |||
Db      241  ENVYVMKQVVDVKAVTSNVASVQCDNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300
      |||
Qy      301  EGDKVRPHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 360
      |||
Db      301  EGDKVRPHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 360
      |||
Qy      361  WYVDGVEVHNAKTPREQYNSYTRVVSVLTVLHQDMLNGKEYKCKVSNKALPVPLEKTI 420
      |||
Db      361  WYVDGVEVHNAKTPREQYNSYTRVVSVLTVLHQDMLNGKEYKCKVSNKALPVPLEKTI 420
      |||
Qy      421  SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 480
      |||
Db      421  SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 480
      |||
Qy      481  VLDSGSEFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTKQSLSPGK 531
      |||
Db      481  VLDSGSEFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTKQSLSPGK 531
      |||

```

## RESULT 3

```

US-10-068-426-6
; Sequence 6, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platiect Glycoprotein IB Alpha Fusion Polypeptides and
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(531)
; OTHER INFORMATION: GPIb290/1A-IG
US-10-068-426-6

```

Query Match 99.4%; Score 2823; DB 15; Length 531;  
 Best Local Similarity 99.4%; Pred. No. 7.5e-218;  
 Matches 528; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

Qy      1  MPLLLLLLLLPSPLPHPICEVSKVASHLEVNCNKNLTALPPDLPKDITILHSENLY 60
      |||
Db      1  MPLLLLLLLLPSPLPHPICEVSKVASHLEVNCNKNLTALPPDLPKDITILHSENLY 60
      |||
Qy      61  TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLTGLDLSHNLQSLPLLGOTLPALT 120
      |||
Db      61  TFSLATLMPYTRLTQNLDRCELTKLQVDGTLPLVLTGLDLSHNLQSLPLLGOTLPALT 120
      |||
Qy      121  LDVSFNRUTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTSLP 180
      |||
Db      121  LDVSFNRUTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTSLP 180
      |||

```

Qy	181	AGLLNGLENLDTLLLOENSLYTI	PKGFTGSHLLPFAFLHGNPML	CNCEILYFRRWLQDNA	240
Db	181	AGLLNGLENLDTLLLOENSLYTI	PKGFTGSHLLPFAFLHGNPML	CNCEILYFRRWLQDNA	240
Qy	241	ENVYVMKQVVDVAVTSNVS	VQCDNSDKFPVYK	PKGCGPTLGDGDTLDYDYPEEDT	300
Db	241	ENVYVMKQVVDVAVMTSNVS	VQCDNSDKFPVYK	PKGCGPTLGDGDTLDYDYPEEDT	300
Qy	301	EGDKVRPHTCPPCPAPEALGAPSV	FLFPKPKDTLMISRTPEVTCVVVDV	SHEDPEVKFN	360
Db	301	EGDKVRPHTCPPCPAPEALGAPSV	FLFPKPKDTLMISRTPEVTCVVVDV	SHEDPEVKFN	360
Qy	361	WYVDGVEVHNATKPREQVNSTYRW	SVLTVLHQDWLNGKEYKCKVSNKALP	VPVPIEKTI	420
Db	361	WYVDGVEVHNATKPREQVNSTYRW	SVLTVLHQDWLNGKEYKCKVSNKALP	VPVPIEKTI	420
Qy	421	SKAGQPREPQVYTLPPSREEMTKNQ	VSLSLTCLVKGFFPSDIAVEWESNGQP	ENNYKTTTP	480
Db	421	SKAGQPREPQVYTLPPSREEMTKNQ	VSLSLTCLVKGFFPSDIAVEWESNGQP	ENNYKTTTP	480
Qy	481	VLDSDGSFFLYSKLTVDKSRQOQGNV	PSCSVMHEALHNHYTQKSLSLSPGK	531	
Db	481	VLDSDGSFFLYSKLTVDKSRQOQGNV	PSCSVMHEALHNHYTQKSLSLSPGK	531	

```

RESULT 4
US-10-068-426-3
; Sequence 3, Application US/10068426
; Publication NO. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; FILE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/286,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(544)
; OTHER INFORMATION: GPlb302/4X-Ig
US-10-068-426-3

```

Db	181	AGLLNGLENLDTLLLQRLNSLTYTPKGFFGGSHLLPFAFLHGNPWLNCCEILYFRFRWLQDNA	241
Qy	241	ENVYVMKQVVDVKATSNVASVQCDNSDKFPVYKYPGKGCPPTLGDDEGDTLDLYDYYPEEDT	300
Db	241	ENVYVMKQVVDVKATSNVASVQCDNSDKFPVYKYPGKGCPPTLGDDEGDTLDLYDYYPEEDT	300
Qy	301	EGDKV-----RPHTCPPCPAPALGAPSVFLPPPKPKOTLMISRTPEVTCVV	347
Db	301	EGDKVAATATVYKEPTKARPHTCPPCPAPALGAPSVFLPPPKPKOTLMISRTPEVTCVV	360
Qy	348	VDVSHEDPEVKFNMTVDGVEVHNAKTPREEQYNSTYRVVSLTVLHQDWLNGKEYCKV	407
Db	361	VDVSHEDPEVKFNMTVDGVEVHNAKTPREEQYNSTYRVVSLTVLHQDWLNGKEYCKV	420
Qy	408	SNKALPVPPIEKTISKAKQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEMES	467
Db	421	SNKALPVPPIEKTISKAKQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEMES	480
Qy	468	NGQPENNYKTTTPVLDSGSEFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSL	527
Db	481	NGQPENNYKTTTPVLDSGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSL	540
Qy	528	SPGK 531	
Db	541	SPGK 544	

RESULT 5  
 US-10-068-426-1  
 ; Sequence 1, Application US/10068426  
 ; Publication No. US20030091576A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shaw, Gray D.  
 ; APPLICANT: Sako, Dianne S.  
 ; APPLICANT: Kumar, Ravindra  
 ; APPLICANT: Sullivan, Francis  
 ; APPLICANT: McDonagh, Tom  
 ; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and  
 ; TITLE OF INVENTION: Methods of Use Thereof  
 ; FILE REFERENCE: 22058-503  
 ; CURRENT APPLICATION NUMBER: US/10/068,426  
 ; CURRENT FILING DATE: 2002-02-06  
 ; PRIOR APPLICATION NUMBER: 60/266,838  
 ; PRIOR FILING DATE: 2001-02-06  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 544  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: DOMAIN  
 ; LOCATION: (1)..(544)  
 ; OTHER INFORMATION: GPlb302-Ig  
 US-10-068-426-1

Db 181 AGLNGLENLDTLLQENSLYTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRMLQDNA 240  
Qy 241 ENVYWKQVVDKAVTSNVASVQCNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300  
Db 241 ENVYWKQVVDKAVTSNVASVQCNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300  
Qy 301 EGDKV-----RPHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 347  
Db 301 EGDKVRAITVVVFTKARPHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 360  
Qy 348 VDVSHEDEPVKFNWYVDGVEVHNATKPREQYNSYRVSVLTVLHODMLNGKEYCKV 407  
Db 361 VDVSHEDEPVKFNWYVDGVEVHNATKPREQYNSYRVSVLTVLHODMLNGKEYCKV 420  
Qy 408 SNKALPVPLEKTSKAKQPREPOVYTLPPSREMTKNQVSLTCLVKGFPSPDIAVEMES 467  
Db 421 SNKALPVPLEKTSKAKQPREPOVYTLPPSREMTKNQVSLTCLVKGFPSPDIAVEMES 480  
Qy 468 NGOPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTKQSL 527  
Db 481 NGOPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTKQSL 540  
Qy 528 SPKG 531  
Db 541 SPKG 544

RESULT 6  
US-10-068-426-2  
; Sequence 2, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaw, Gray D.  
; APPLICANT: Sako, Dianne S.  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Sullivan, Francis  
; APPLICANT: McDonagh, Tom  
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and  
; FILE OF INVENTION: Methods of Use Thereof  
; CURRENT APPLICATION NUMBER: US/10/068,426  
; CURRENT FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: 60/266,838  
; PRIOR FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 544  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(544)  
; OTHER INFORMATION: GP1b302/2A-Ig  
US-10-068-426-2

Query Match 99.1%; Score 2812.5; DB 15; Length 544;  
Best Local Similarity 97.2%; Pred. No. 5.4e-217;  
Matches 529; Conservative 1; Mismatches 1; Indels 13; Gaps 1;  
Qy 1 MPLLLLLLLLPSLPHPICEVSKVASHLEVNCCKRNLTAIPDLPKDITLHLSNLLY 60  
Db 1 MPLLLLLLLLPSLPHPICEVSKVASHLEVNCCKRNLTAIPDLPKDITLHLSNLLY 60  
Qy 61 TFSLATLAPYRLTQNLDRCELTKLQVDTLPVLTGLDLSHNQIQLGQTLPLATY 120  
Db 61 TFSLATLAPYRLTQNLDRCELTKLQVDTLPVLTGLDLSHNQIQLGQTLPLATY 120  
Qy 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTPL 180  
Db 121 LDVSNRLTSLPLGALRGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTPL 180

Qy 181 AGLNGLENLDTLLQENSLYTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRMLQDNA 240  
Db 181 AGLNGLENLDTLLQENSLYTI PKGFFGSHLLPFAFLHGNPWLNCCEILYFRMLQDNA 240  
Qy 241 ENVYWKQVVDKAVTSNVASVQCNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300  
Db 241 ENVYWKQVVDKAVTSNVASVQCNSDKFPYKYPGKGCPTLGDGDTLDLYYPEEDT 300  
Qy 301 EGDKV-----RPHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 347  
Db 301 EGDKVAATAVVVFTKARPHTCPCPAPEALGAPSVFLFPKPKDTLMISRTPEVTCV 360  
Qy 348 VDVSHEDEPVKFNWYVDGVEVHNATKPREQYNSYRVSVLTVLHODMLNGKEYCKV 407  
Db 361 VDVSHEDEPVKFNWYVDGVEVHNATKPREQYNSYRVSVLTVLHODMLNGKEYCKV 420  
Qy 408 SNKALPVPLEKTSKAKQPREPOVYTLPPSREMTKNQVSLTCLVKGFPSPDIAVEMES 467  
Db 421 SNKALPVPLEKTSKAKQPREPOVYTLPPSREMTKNQVSLTCLVKGFPSPDIAVEMES 480  
Qy 468 NGOPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTKQSL 527  
Db 481 NGOPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTKQSL 540  
Qy 528 SPKG 531  
Db 541 SPKG 544

RESULT 7  
US-10-068-426-11  
; Sequence 11, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:  
; APPLICANT: Shaw, Gray D.  
; APPLICANT: Sako, Dianne S.  
; APPLICANT: Kumar, Ravindra  
; APPLICANT: Sullivan, Francis  
; APPLICANT: McDonagh, Tom  
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and  
; FILE OF INVENTION: Methods of Use Thereof  
; CURRENT APPLICATION NUMBER: US/10/068,426  
; CURRENT FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: 60/266,838  
; PRIOR FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 290  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(290)  
; OTHER INFORMATION: GB1b290/2V  
US-10-068-426-11

Query Match 54.1%; Score 1536; DB 15; Length 290;  
Best Local Similarity 100.0%; Pred. No. 4.8e-115;  
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 17 HPICEVSKVASHLEVNCCKRNLTAIPDLPKDITLHLSNLLYFSLATLMPYRLTQ 76  
Db 1 HPICEVSKVASHLEVNCCKRNLTAIPDLPKDITLHLSNLLYFSLATLMPYRLTQ 60  
Qy 77 NDRCSLTKLQVDTLPVLTGLDLSHNQIQLGQTLPLATYLDVSNRLTSLPLGAL 136  
Db 61 NDRCSLTKLQVDTLPVLTGLDLSHNQIQLGQTLPLATYLDVSNRLTSLPLGAL 120  
Qy 137 RGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTPLPAGLNGLENLDTLLQ 196  
Db 121 RGELQELYLKGNELKTLPPGLLTPTPKLEKLSANNLTPLPAGLNGLENLDTLLQ 180

Qy	197	ENSLYTI	PKGFGSHLLPFAFLHGNPMWLCNCEILLYFRRLWQDNNAENTVWKQVVDVKAVT	256
Db	181	ENSLYTI	PKGFGSHLLPFAFLHGNPMWLCNCEILLYFRRLWQDNNAENTVWKQVVDVKAVT	240
Qy	257	SNVASQC	DNSDKFVYVYKPGKCGPTLGDGEGDTLDLYYFSEEDTEGDKVR	306
Db	241	SNVASQC	DNSDKFVYVYKPGKCGPTLGDGEGDTLDLYYFSEEDTEGDKVR	290

## RESULT 8

```

US-10-068-426-7
; Sequence 7, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(302)
; OTHER INFORMATION: GPlb302
US-10-068-426-7

```

## RESIT.T 9

US-10-068-426-10  
; Sequence 10, Application US/10068426  
; Publication No. US20030091576A1  
; GENERAL INFORMATION:



```

; NAME/KEY: DOMAIN
; LOCATION: (1)..(302)
; OTHER INFORMATION: GP1b302/2A
US-10-068-426-8

Query Match      53.6%; Score 1521.5; DB 15; Length 302;
Best Local Similarity 92.7%; Pred. No. 7.4e-114;
Matches 291; Conservative 2; Mismatches 8; Indels 13; Gaps 1;

QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTHLHSENLLYTFSLATLMPYTR:TQL 76
DB 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTHLHSENLLYTFSLATLMPYTR:TQL 60
QY 77 NLDRCELTKLQVDGTLPLVGLTDLDSHNOQLSLPLGOTLPALTVDVSFNRLTSLPLGAL 136
DB 61 NLDRCELTKLQVDGTLPLVGLTDLDSHNOQLSLPLGOTLPALTVDVSFNRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196
DB 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180
QY 197 ENSLYTIIPKGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNAENVYVWKQVVDVKAVT 256
DB 181 ENSLYTIIPKGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNAENVYVWKQVVDVKAVT 240
QY 257 SNVASVQCDNSDKFPVYKPGKGPCTLGDGDTLDYDYPEEDTEGDKVR 306
DB 241 SNVASVQCDNSDKFPVYKPGKGPCTLGDGDTLDYDYPEEDTEGDKVR 290

RESULT 12
US-10-068-426-9
; Sequence 9, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 9
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(301)
; OTHER INFORMATION: GP1b/4X
US-10-068-426-9

Query Match      53.3%; Score 1514; DB 15; Length 301;
Best Local Similarity 93.0%; Pred. No. 2.9e-113;
Matches 292; Conservative 1; Mismatches 7; Indels 14; Gaps 2;

QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTHLHSENLLYTFSLATLMPYTR:TQL 76
DB 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTHLHSENLLYTFSLATLMPYTR:TQL 60
QY 77 NLDRCELTKLQVDGTLPLVGLTDLDSHNOQLSLPLGOTLPALTVDVSFNRLTSLPLGAL 136
DB 61 NLDRCELTKLQVDGTLPLVGLTDLDSHNOQLSLPLGOTLPALTVDVSFNRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196
DB 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180
QY 197 ENSLYTIIPKGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNAENVYVWKQVVDVKAVT 256
DB 181 ENSLYTIIPKGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNAENVYVWKQVVDVKAVT 240
QY 257 SNVASVQCDNSDKFPVYKPGKGPCTLGDGDTLDYDYPEEDTEGDKVRPHTCPPCAP 316
DB 241 SNVASVQCDNSDKFPVYKPGKGPCTLGDGDTLDYDYPEEDTEGDKV----- 288
QY 317 EALGAPSVFLPPK 330
DB 289 --AATATVVKPPTK 300

; NAME/KEY: DOMAIN
; LOCATION: (1)..(302)
; OTHER INFORMATION: GP1b302/2A
US-10-068-426-8

Query Match      53.6%; Score 1521.5; DB 15; Length 302;
Best Local Similarity 92.7%; Pred. No. 7.4e-114;
Matches 291; Conservative 2; Mismatches 8; Indels 13; Gaps 1;

QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTHLHSENLLYTFSLATLMPYTR:TQL 76
DB 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTHLHSENLLYTFSLATLMPYTR:TQL 60
QY 77 NLDRCELTKLQVDGTLPLVGLTDLDSHNOQLSLPLGOTLPALTVDVSFNRLTSLPLGAL 136
DB 61 NLDRCELTKLQVDGTLPLVGLTDLDSHNOQLSLPLGOTLPALTVDVSFNRLTSLPLGAL 120
QY 137 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196
DB 121 RGLGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180
QY 197 ENSLYTIIPKGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNAENVYVWKQVVDVKAVT 256
DB 181 ENSLYTIIPKGFPGSHLLPFAFLHGNPWLNCCEILYFRWLQDNAENVYVWKQVVDVKAVT 240
QY 257 SNVASVQCDNSDKFPVYKPGKGPCTLGDGDTLDYDYPEEDTEGDKVRPHTCPPCAP 316
DB 241 SNVASVQCDNSDKFPVYKPGKGPCTLGDGDTLDYDYPEEDTEGDKV----- 289
QY 317 EALGAPSVFLPPK 330
DB 290 --AATATVVKPPTK 301

RESULT 11
US-10-068-426-12
; Sequence 12, Application US/10068426
; Publication No. US20030091576A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Gray D.
; APPLICANT: Sako, Dianne S.
; APPLICANT: Kumar, Ravindra
; APPLICANT: Sullivan, Francis
; APPLICANT: McDonagh, Tom
; TITLE OF INVENTION: Platelet Glycoprotein IB Alpha Fusion Polypeptides and
; TITLE OF INVENTION: Methods of Use Thereof
; FILE REFERENCE: 22058-503
; CURRENT APPLICATION NUMBER: US/10/068,426
; CURRENT FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/266,838
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 12
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(290)
; OTHER INFORMATION: GB1b290/1A
US-10-068-426-12

Query Match      53.5%; Score 1520; DB 15; Length 290;
Best Local Similarity 99.0%; Pred. No. 9.2e-114;
Matches 287; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTHLHSENLLYTFSLATLMPYTR:TQL 76
DB 1 HPICEVSKVASHLEVNCCKRNLTALPPDLPKDTHLHSENLLYTFSLATLMPYTR:TQL 60
QY 77 NLDRCELTKLQVDGTLPLVGLTDLDSHNOQLSLPLGOTLPALTVDVSFNRLTSLPLGAL 136

```

RESULT 13

US-09-825-580-4  
; Sequence 4, Application US/09825580  
; Publication No. US20030166521A1  
; GENERAL INFORMATION:  
; APPLICANT: Eppihimer, Michael J.  
; APPLICANT: Schaub, Robert G.  
; APPLICANT: Harris, Alan  
; TITLE OF INVENTION: Inhibition of Thrombosis by Treatment with  
; FILE REFERENCE: GFN-5398  
; CURRENT APPLICATION NUMBER: US/09/825,580  
; CURRENT FILING DATE: 2001-04-02  
; PRIOR APPLICATION NUMBER: 60/193,787  
; PRIOR FILING DATE: 2000-03-31  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 313  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-825-580-4

Query Match 43.6%; Score 1237; DB 12; Length 313;  
Best Local Similarity 97.0%; Pred. No. 5e-91;  
Matches 229; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
  
QY 296 PEEDTEGDKVRPHCTPCCPAPEALGAPSVFLPPKPKDTLMISRTPVTCVVVDVSHEDP 355  
DB 78 PESTTVEPAARPHCTPCCPAPEALGAPSVFLPPKPKDTLMISRTPVTCVVVDVSHEDP 137  
  
QY 356 EVKFNMYVDGVEVHNKTKPRBEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVP 415  
DB 138 EVKFNMYVDGVEVHNKTKPRBEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVP 197  
  
QY 416 IKTISKAKGQPREPQVYTLPPSRREMTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNY 475  
DB 198 IKTISKAKGQPREPQVYTLPPSRREMTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNY 257  
  
QY 476 KTTTPPVLDSDGSFPLYSKLTVDKSRWQOQGNVFSCSVMEALHNHYTKLSLSLSPGK 531  
DB 258 KTTTPPVLDSDGSFPLYSKLTVDKSRWQOQGNVFSCSVMEALHNHYTKLSLSLSPGK 313

RESULT 14

US-10-211-786-4  
; Sequence 4, Application US/10211786  
; Publication No. US20030083258A1  
; GENERAL INFORMATION:  
; APPLICANT: Michael J. Eppihimer  
; APPLICANT: Robert G. Schaub  
; APPLICANT: Ronald Tuma  
; TITLE OF INVENTION: MODULATION OF LEUKOCYTE-ENDOTHELIAL INTERACTIONS FOLLOWING ISCHEM  
; FILE REFERENCE: 8702.0099-00000  
; CURRENT APPLICATION NUMBER: US/10/211,786  
; CURRENT FILING DATE: 2002-08-02  
; PRIOR APPLICATION NUMBER: 60/309,816  
; PRIOR FILING DATE: 2001-08-03  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 313  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-211-786-4

Query Match 43.6%; Score 1237; DB 15; Length 313;  
Best Local Similarity 97.0%; Pred. No. 5e-91;  
Matches 229; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
  
QY 296 PEEDTEGDKVRPHCTPCCPAPEALGAPSVFLPPKPKDTLMISRTPVTCVVVDVSHEDP 355  
DB 78 PESTTVEPAARPHCTPCCPAPEALGAPSVFLPPKPKDTLMISRTPVTCVVVDVSHEDP 137

QY 356 EVKFNMYVDGVEVHNKTKPRBEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVP 415  
DB 138 EVKFNMYVDGVEVHNKTKPRBEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPVP 197  
  
QY 416 IKTISKAKGQPREPQVYTLPPSRREMTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNY 475  
DB 198 IKTISKAKGQPREPQVYTLPPSRREMTKNQVSLTCLVKGFPYSDIAVEWESNGQPENNY 257  
  
QY 476 KTTTPPVLDSDGSFPLYSKLTVDKSRWQOQGNVFSCSVMEALHNHYTKLSLSLSPGK 531  
DB 258 KTTTPPVLDSDGSFPLYSKLTVDKSRWQOQGNVFSCSVMEALHNHYTKLSLSLSPGK 313

RESULT 15  
US-10-264-634-33  
; Sequence 33, Application US/10264634  
; Publication No. US20030108549A1  
; GENERAL INFORMATION:  
; APPLICANT: Donaldson, Debra et al.  
; TITLE OF INVENTION: Methods and Compositions for Modulating Interleukin-21 Receptor;  
; FILE REFERENCE: GI5320-P3  
; CURRENT APPLICATION NUMBER: US/10/264,634  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: 09/040,005  
; PRIOR FILING DATE: 1998-03-17  
; PRIOR APPLICATION NUMBER: 09/560,766  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 09/569,384  
; PRIOR FILING DATE: 2000-05-11  
; PRIOR APPLICATION NUMBER: 09/972,218  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: 60/373,746  
; PRIOR FILING DATE: 2002-04-17  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: Patent In version 3.1  
; SEQ ID NO 33  
; LENGTH: 470  
; TYPE: PRT  
; ORGANISM: Human  
US-10-264-634-33

Query Match 43.4%; Score 1232.5; DB 15; Length 470;  
Best Local Similarity 74.6%; Pred. No. 2e-90;  
Matches 241; Conservative 16; Mismatches 39; Indels 27; Gaps 3;  
  
QY 220 GNPWLCNCEILYFRWLQDNANVYVWKQVVDKAVTSNVASVQCDNSDKFPVYKYPGKG 279  
DB 164 GDPMAVSP-----RRKLIS-----VDSRSVSLLEFRKQSSYELQVRAGMPG 207  
  
QY 280 C-----PTLGDGEGDTLDVYYPBEDTEGDKVRPHCTPCCPAPEALGAPSVFLFP 328  
DB 208 SSYQGTWSESDPVIFOTQSEELKEGWNDDDDKSGDKTHTCCPCPAPEALGAPSVFLFP 267  
  
QY 329 PKPKDTLMISRTPVTCVVVDVSHEDPEVKFNMYVDGVEVHNKTKPRBEQYNSTYRVVS 388  
DB 268 PKPKDTLMISRTPVTCVVVDVSHEDPEVKFNMYVDGVEVHNKTKPRBEQYNSTYRVVS 327  
  
QY 389 VLTVLHQDWLNGKEYKCKVSNKALPVPPIETKISKAKGQPREPQVYTLPPSRREMTKNQVS 448  
DB 328 VLTVLHQDWLNGKEYKCKVSNKALPVPPIETKISKAKGQPREPQVYTLPPSRREMTKNQVS 387  
  
QY 449 LTCLVKGFPYSDIAVEWESNGQPENNYKTTTPPVLDSDGSFPLYSKLTVDKSRWQOQGNVES 508  
DB 388 LTCLVKGFPYSDIAVEWESNGQPENNYKTTTPPVLDSDGSFPLYSKLTVDKSRWQOQGNVES 447

Search completed: October 14, 2003, 06:48:24  
Job time : 195.642 secs

-----

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: October 14, 2003, 06:17:26 ; Search time 21.734 Seconds  
(without alignments)  
1033.731 Million call updates/sec

Title: US-10-068-426-5

Perfect score: 2839

Sequence: 1 MPLLILLLLPLPHPIC.....MHEALHNHYTKLSLSPGK 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first: 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*

2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*

4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*

5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*

6: /cgn2\_6/ptodata/2/1aa/backfiles.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1528.5	53.8	320	1	US-07-613-083B-1
2	1521.5	53.6	610	1	US-07-821-717B-6
3	1521.5	53.6	610	1	US-08-119-262B-6
4	1521.5	53.6	610	1	US-08-135-929A-11
5	1521.5	53.6	610	1	US-08-234-265A-11
6	1237	43.6	313	3	US-08-713-556F-36
7	1230	43.3	388	3	US-09-131-247-16
8	1230	43.3	449	3	US-08-897-236-23
9	1230	43.3	449	3	US-09-500-253B-23
10	1223	43.1	451	2	US-08-887-352B-14
11	1223	43.1	451	2	US-08-887-352B-16
12	1223	43.1	451	2	US-08-887-352B-18
13	1223	43.1	451	3	US-08-466-151-65
14	1223	43.1	451	3	US-09-109-207C-14
15	1223	43.1	451	3	US-09-109-207C-16
16	1223	43.1	451	3	US-09-109-207C-18
17	1223	43.1	451	3	US-09-282-505-2
18	1223	43.1	451	3	US-09-054-255-2
19	1223	43.1	451	3	US-09-296-005-14
20	1223	43.1	451	3	US-09-296-005-16
21	1223	43.1	451	3	US-09-296-005-18
22	1223	43.1	451	4	US-09-282-846-2
23	1223	43.1	451	4	US-09-680-145-2
24	1222	43.0	453	3	US-08-466-151-8
25	1222	43.0	453	4	US-08-466-163B-8
26	1222	43.0	977	4	US-09-590-656-1
27	1222	43.0	977	4	US-09-733-764-1

Sequence 18, Appl  
Sequence 32, Appl  
Sequence 30, Appl  
Sequence 3, Appl  
Sequence 3, Appl  
Sequence 71, Appl  
Sequence 71, Appl  
Sequence 71, Appl  
Sequence 71, Appl  
Sequence 15, Appl  
Sequence 9, Appl  
Sequence 26, Appl  
Sequence 24, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 8, Appl

## ALIGNMENTS

RESULT 1  
US-07-613-083B-1  
; Sequence 1, Application US/07613083B  
; Patent No. 5340727  
; GENERAL INFORMATION:  
; APPLICANT: Ruggeri, Zaverio M.  
; APPLICANT: Ware, Jerry, inventors  
; APPLICANT: on behalf of Scripps Clinic and Research  
; APPLICANT: Foundation  
; TITLE OF INVENTION: GPIb' Fragments and Recombinant  
; TITLE OF INVENTION: DNA Expression Vectors  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scripps Clinic and Research  
; ADDRESSEE: Foundation  
; STREET: 10666 No. 5340727th Torrey Pines Road  
; CITY: La Jolla  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette-5.25 inch, 1.2 Mb  
; COMPUTER: AST Bravo IBM PC comp. (386SX)  
; OPERATING SYSTEM: MS DOS version 3.2  
; SOFTWARE: Wordperfect 5.1 conv. to ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/613,083B  
; FILING DATE: 19911114  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA: This appl. is a c-i-p of  
; APPLICATION NUMBER: U.S. 07/470,674  
; FILING DATE: 04-Jan-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Barron, Alexis  
; REGISTRATION NUMBER: 22,702  
; REFERENCE/DOCKET NUMBER: P16,569-B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 923-4466  
; TELEFAX: (215) 923-2189  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 320  
; TYPE: AMINO ACID  
; STRANDEDNESS: No. 5340727 applicable  
; TOPOLOGY: Linear  
US-07-613-083B-1

Query Match 53.8%; Score 1528.5; DB 1; Length 320;  
Best Local Similarity 95.1%; Pred. No. 2.3e-129;  
Matches 291; Conservative 1; Mismatches 3; Indels 11; Gaps 1;

17 HPICEVSKVASHLEVNCNKRNTALPPDLPKDTTILHLSNLTFTSLATLMPYTRLTOL 76  
1 HPICEVSKVASHLEVNCNKRNTALPPDLPKDTTILHLSNLTFTSLATLMPYTRLTOL 60  
77 NLDRCETLKQVDTLPVLTGLDLSHNQOSLPGLQTLPAITLVDSFNRLTSLPLGAL 136  
61 NLDRCETLKQVDTLPVLTGLDLSHNQOSLPGLQTLPAITLVDSFNRLTSLPLGAL 120  
137 RGLGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196  
121 RGLGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180  
197 ENSLYTIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLQDQNAENVYWKQVDVKAVT 256  
181 ENSLYTIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLQDQNAENVYWKQVDVKAVT 240  
257 SNVASVQCDNSDKFPVYKPGKCGPTLGDEGDTLDYDYYPEEDTEGDKVYR----- 306  
241 SNVASVQCDNSDKFPVYKPGKCGPTLGDEGDTLDYDYYPEEDTEGDKVYR----- 300  
307 -PHTCP 311  
301 KAHTTP 306

RESULT 2  
US-07-821-717B-6  
; Sequence 6, Application: US/07821717B  
; Patent No. 5298239  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Cunningham, David  
; APPLICANT: Lyle, Vicki A.  
; APPLICANT: Finch, Clara N.  
; TITLE OF INVENTION: MUTATIONS RENDERING PLATELET  
; TITLE OF INVENTION: GLYCOPROTEIN IB ALPHA LESS REACTIVE  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/821,717B  
; FILING DATE: 15-JAN-1992  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Timain, Susan J.  
; REGISTRATION NUMBER: 34,103  
; REFERENCE/DOCKET NUMBER: 20884/21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1636  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PUBLICATION INFORMATION:  
; AUTHORS: Lopez, Jose A.  
; AUTHORS: Chung, Dominic W.  
; AUTHORS: Fujikawa, Kazuo  
; AUTHORS: Hagen, Frederick S.

AUTHORS: Papayannopoulou, Thalia  
AUTHORS: Roth, Gerald J.  
TITLE: Cloning of the alpha chain of human  
TITLE: platelet glycoprotein Ib: A transmembrane protein with homology  
TITLE: to leucine-rich alpha-2-glycoprotein  
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
VOLUME: 84  
PAGES: 5615-5619  
DATE: AUG-1987  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 610  
PUBLICATION INFORMATION:  
AUTHORS: Zimmerman, Theodore S.  
AUTHORS: Ruggeri, Zaverio M.  
AUTHORS: Houghten, Richard A.  
AUTHORS: Vincete, Vincete  
AUTHORS: Mohri, Hiroshi  
TITLE: Proteolytic fragments and synthetic  
TITLE: peptides that block the binding of von Willebrand factor to the  
TITLE: platelet membrane glycoprotein Ib  
DOCUMENT NUMBER: EP 0 317 278 A2  
FILING DATE: 16-NOV-1988  
PUBLICATION DATE: 24-MAY-1989  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 293  
US-07-821-717B-6  
Query Match 53.6%; Score 1521.5; DB 1; Length 610;  
Best Local Similarity 82.9%; Pred. No. 2.6e-128;  
Matches 301; Conservative 5; Mismatches 22; Indels 35; Gaps 5;  
QY 17 HPICEVSKVASHLEVNCNKRNTALPPDLPKDTTILHLSNLTFTSLATLMPYTRLTOL 76  
DB 1 HPICEVSKVASHLEVNCNKRNTALPPDLPKDTTILHLSNLTFTSLATLMPYTRLTOL 60  
QY 77 NLDRCETLKQVDTLPVLTGLDLSHNQOSLPGLQTLPAITLVDSFNRLTSLPLGAL 136  
DB 61 NLDRCETLKQVDTLPVLTGLDLSHNQOSLPGLQTLPAITLVDSFNRLTSLPLGAL 120  
QY 137 RGLGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTLPAGLNGLENLDTLLQ 196  
DB 121 RGLGELQELYLKGNELKTLPPGLTTPPKLEKLSLANNLTLPAGLNGLENLDTLLQ 180  
QY 197 ENSLYTIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLQDQNAENVYWKQVDVKAVT 256  
DB 181 ENSLYTIPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLQDQNAENVYWKQVDVKAVT 240  
QY 257 SNVASVQCDNSDKFPVYKPGKCGPTLGDEGDTLDYDYYPEEDTEGDKVYR----- 306  
DB 241 SNVASVQCDNSDKFPVYKPGKCGPTLGDEGDTLDYDYYPEEDTEGDKVYR----- 300  
QY 307 -PHTCP-----PCPA---PEALGAPSVFLFPPK---PKDTL-----MISRT 341  
DB 301 KAHTTPWGLFYSWSTASLDSOMPSSLHPTQESTKEQITFFPRWTFNFTLHMSITFSKTP 360  
QY 342 EVT 344  
DB 361 KST 363  
RESULT 3  
US-08-119-262B-6  
; Sequence 6, Application: US/08119262B  
; Patent No. 5492809  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; APPLICANT: Cunningham, David  
; APPLICANT: Lyle, Vicki A.  
; APPLICANT: Finch, Clara N.  
; TITLE OF INVENTION: MUTATIONS RENDERING PLATELET  
; TITLE OF INVENTION: GLYCOPROTEIN IB ALPHA LESS REACTIVE  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051

CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM: disk  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/119,262B  
FILING DATE: 09-SEP-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/821,717  
FILING DATE: 15-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Timian, Susan J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 20884/22  
TELEPHONE: (716) 263-1636  
TELEFAX: (716) 263-1600  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PUBLICATION INFORMATION:  
AUTHORS: Lopez, Jose A.  
AUTHORS: Chung, Dominic W.  
AUTHORS: Fujikawa, Kazuo  
AUTHORS: Hagen, Frederick S.  
AUTHORS: Papayannopoulou, Thalia  
AUTHORS: Roth, Gerald J.  
TITLE: Cloning of the alpha chain of human platelet glycoprotein Ib: A transmembrane protein  
TITLE: leucine-rich alpha-2-glycoprotein  
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
VOLUME: 84  
PAGES: 5615-5619  
DATE: AUG-1987  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 610  
PUBLICATION INFORMATION:  
AUTHORS: Zimmerman, Theodore S.  
AUTHORS: Ruggeri, Zaverio M.  
AUTHORS: Houghten, Richard A.  
AUTHORS: Vincete, Vincete  
AUTHORS: Mohri, Hiroshi  
TITLE: Proteolytic fragments and synthetic peptides that block the binding of von Willebrand factor to the platelet glycoprotein Ib  
TITLE: membrane glycoprotein Ib  
DOCUMENT NUMBER: EP 0 317 278 A2  
FILING DATE: 16-NOV-1988  
PUBLICATION DATE: 24-MAY-1989  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 293  
US-08-119-262B-6  
Query Match 53.6%; Score 1521.5; DB 1; Length 610;  
Best Local Similarity 82.9%; Pred. No. 2.6e-128;  
Matches 301; Conservative 5; Mismatches 22; Indels 35; Gaps 5;  
QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLKPTTLHLSNLLYTFSLATLMPYTRLTOL 76  
DB 1 HPICEVSKVASHLEVNCCKRNLTALPPDLKPTTLHLSNLLYTFSLATLMPYTRLTOL 60  
QY 77 NLDRCCLTKLQVDTLPLVGLTDLSHNQLQSLPLGLQTLPALTVLDVSNFRLTSLPLGAL 136  
DB 61 NLDRCCLTKLQVDTLPLVGLTDLSHNQLQSLPLGLQTLPALTVLDVSNFRLTSLPLGAL 120  
QY 137 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLPVGLNGLENLDTLLQ 196

DB 121 RGLGELQELYLKGNELKTLPPGLLTPPKLEKLSLANNLTLPVGLNGLENLDTLLQ 180  
QY 197 ENSLYTTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLWLDNAENYVVKQVVDKAVT 256  
DB 181 ENSLYTTPKGFSGSHLLPFAFLHGNPWLNCCEILYFRRLWLDNAENYVVKQVVDKAVT 240  
QY 257 SNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDLYDYDEEDTEGDKVR----- 306  
DB 241 SNVASVQCDNSDKFPVYKPGKGCPTLGDGDTLDLYDYDEEDTEGDKVRATRVVKFPT 300  
QY 307 -PHTCP-----PCPA---PEALGAPSVFLFPK--PKDTL-----MISRTTP 341  
DB 301 KAHTTPWGLFYVSWSTASLDSQMPSSLHPTQESTKEQTTPFPRTNPTLHMSITFSKTP 360  
QY 342 EVT 344  
DB 361 KST 363

RESULT 4  
US-08-135-929A-11  
Sequence 11, Application US/08135929A  
Patent No. 5593959  
GENERAL INFORMATION:  
APPLICANT: Miller, Jonathan L.  
APPLICANT: Cunningham, David  
APPLICANT: Lyle, Vicki A.  
APPLICANT: Finch, Clara N.  
APPLICANT: Pincus, Matthew R.  
TITLE OF INVENTION: Mutations in the Gene Encoding the Alpha  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/135,929A  
FILING DATE: 14-OCT-1993  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Timian, Susan J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 20884/23  
TELEPHONE: (716) 263-1636  
TELEFAX: (716) 263-1600  
TELEX: 978450  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-135-929A-11  
Query Match 53.6%; Score 1521.5; DB 1; Length 610;  
Best Local Similarity 82.9%; Pred. No. 2.6e-128;  
Matches 301; Conservative 5; Mismatches 22; Indels 35; Gaps 5;  
QY 17 HPICEVSKVASHLEVNCCKRNLTALPPDLKPTTLHLSNLLYTFSLATLMPYTRLTOL 76  
DB 1 HPICEVSKVASHLEVNCCKRNLTALPPDLKPTTLHLSNLLYTFSLATLMPYTRLTOL 60

[illegible]

```

RESULT 5
US-08-234-265A-11
; Sequence 11, Application US/08234265A
; Patent NO. 5624817
; GENERAL INFORMATION:
; APPLICANT: Miller, Jonathan L.
; APPLICANT: Cunningham, David
; APPLICANT: Lyle, Vicki A.
; APPLICANT: Finch, Clara N.
; APPLICANT: Fincus, Matthew R.
; TITLE OF INVENTION: Mutations in the Gene Encoding the
; TITLE OF INVENTION: Chain of Platelet Glycoprotein Ib
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/234,265A
; FILING DATE: 28-APR-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 20884/24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600
; TELEX: 978450
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-234-265A-11

```

Query Match 53.6%; Score 1521.5; DB 1; Length 610;  
Best Local Similarity 82.9%; Pred. No. 2.6e-128;  
Matches 301; Conservative 5; Mismatches 22; Indels 35; Gaps 5;

Qy	17	HPICEVSKVASHLEWNCQRNLTALPPDLPKDPTTILHLSENLYTSLATIMPYTRLTQ	76
Db	1	HPICEVSKVASHLEWNCQRNLTALPPDLPKDPTTILHLSENLYTSLATIMPYTRLTQ	60
Qy	77	NLDRCELTKI QVDGTPVLVGLTDLISHNQQLSLPUGQTLPALTVDVSNFRLTSLP	136
Db	61	NLDRCELTKI QVDGTPVLVGLTDLISHNQQLSLPUGQTLPALTVDVSNFRLTSLP	120
Qy	137	RGELGELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTETELPAGLLNGLENL	196
Db	121	RGELGELQELYLKAVNELKTLPPGLLTPTPKLEKLSLANNLTETELPVGLLNGLENL	180
Qy	197	ENSLYTI PKPFGGSHLLPFAFLHGNPWL CNCEILYFRWLQD NAEVYVWKQVDYKAT	256
Db	181	ENSLYTI PKPFGGSHLLPFAFLHGNPWL CNCEILYFRWLQD NAEVYVWKQVDYKAT	240
Qy	257	SNVASQCDNSDKFVPYKYCKGKCGPTLGDEGDTLDYYP EEDTEGDKVR	306
Db	241	SNVASQCDNSDKFVPYKYCKGKCGPTLGDEGDTLDYYP EEDTEGDKVRATRVYKFPT	300
Qy	307	-PHTCP-----PCPA---PEALGAPSVFLPPK--PKDTL-----MISRTP	341
Db	301	KAHTTPWGLFYFSWSTASLD SQMPSSLHPTQESTKEQTTFP RPWTPNFTLHWESI	360
Qy	342	EVT	344
		:	
Db	361	XST	363
		:	

RESULT 6  
 US-08-713-556F-36  
 : Sequence 36, Application US/08713556F  
 : Patent No. 6277975  
 : GENERAL INFORMATION:  
 : APPLICANT: Larsen, Glenn  
 : APPLICANT: Sako, Dianne  
 : APPLICANT: Chang, Xiao Jia  
 : APPLICANT: Veldman, Geertuida M.  
 : APPLICANT: Cumming, Dale  
 : APPLICANT: Kumar, Ravindra  
 : APPLICANT: Shaw, Gray  
 : TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN  
 : NUMBER OF SEQUENCES: 45  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: LEGAL AFFAIRS  
 : STREET: 87 CAMBRIDGE PARK DRIVE  
 : CITY: CAMBRIDGE  
 : STATE: MA  
 : COUNTRY: USA  
 : ZIP: 02140  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent In Release #1.0, Version #1.25  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/713,556F  
 : FILING DATE:  
 : CLASSIFICATION: 536  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 07/965,662  
 : FILING DATE: 23-OCT-1992  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/112,608  
 : FILING DATE: 26-AUG-1993  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: PCT/US93/10168  
 : FILING DATE: 22-OCT-1993  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/235,398  
 : FILING DATE: 28-APR-1994  
 : PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/316,305  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/428,734  
FILING DATE: 25-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BROWN, SCOTT A.  
REGISTRATION NUMBER: 32,724  
REFERENCE/DOCKET NUMBER: GI 5213F  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-713-556F-36

Query Match 43.6%; Score 1237; DB 3; Length 313;  
Best Local Similarity 97.0%; Pred. No. 3.8e-103;  
Matches 229; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 296 PBDTEGDKVYPHTCPPEALGAPSVLPFPKPKDTLMISRTPEVTCVVVDVSHEDP 355  
DB 78 PESTTVEPAARHTCPPEALGAPSVLPFPKPKDTLMISRTPEVTCVVVDVSHEDP 137  
QY 356 EVKFNMYVDGVEVHNNAKTPREQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPVP 415  
DB 138 EVKFNMYVDGVEVHNNAKTPREQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPVP 197  
QY 416 IETISKAKGQPEPOVYTLPPSREMTKNQVSLTCLVKGFPSPDIKAVEMESNGQPENNY 475  
DB 198 IETISKAKGQPEPOVYTLPPSREMTKNQVSLTCLVKGFPSPDIKAVEMESNGQPENNY 257  
QY 476 KTTPPVLDSDGSFFLYSLKLTVDKSRQOQGVNFVSCSVMEALHNYTKSLSLSPGK 531  
DB 258 KTTPPVLDSDGSFFLYSLKLTVDKSRQOQGVNFVSCSVMEALHNYTKSLSLSPGK 313

RESULT 7  
US-09-131-247-16  
Sequence 16, Application US/09131247  
Patent No. 6294170  
GENERAL INFORMATION:  
APPLICANT: Boone, Thomas C.  
APPLICANT: Herzhenson, Susan  
APPLICANT: Bevilacqua, Michael P.  
APPLICANT: Collins, David S.  
TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY  
FILE REFERENCE: A-365F  
CURRENT APPLICATION NUMBER: US/09/131,247  
EARLIER FILING DATE: 1998-08-07  
EARLIER APPLICATION NUMBER: 60/055,185  
EARLIER FILING DATE: 1997-08-08  
EARLIER APPLICATION NUMBER: PCT/US 97/02131  
EARLIER FILING DATE: 1997-02-10  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 16  
LENGTH: 388  
TYPE: PRT  
ORGANISM: Human  
US-09-131-247-16  
Query Match 43.3%; Score 1230; DB 3; Length 388;  
Best Local Similarity 64.5%; Pred. No. 2.3e-102;  
Matches 249; Conservative 20; Mismatches 69; Indels 48; Gaps 6;  
QY 171 LANNLTLPAGLNG-----LNNLDTLLQENSLYTIKPGFFGSHLLPFAFLHGNPMLC 225

DB 26 LRNN---QLVAGYLQGNVNLLEKIDVVPIEPHAL-----FLGIRGKMCL 68  
QY 226 NCEILYFRRLQDANAEVYVWKQVVDKAVTSNVASVQCDNSDKFPVYKYPCKGCP-----281  
DB 69 SCVKSQDQETRLQLEAVN-----ITDISENKKQDKRFAFIRSDSGPTTSFESAACPGWFL 122  
QY 282 -----TLGDEGGDTLDYYPPEEDTEGD---KVRPHTCPPCPAPEALGAPSVF 325  
DB 123 CTAMEADQPVSLTNMPDEGVMTKFFQEDAEAAEPKSSDKTHTCTCPAPPELLGGSPV 182  
QY 326 LFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYR 385  
DB 183 LFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYR 242  
QY 386 VVSVLTVLHODWLNKGYCKVSNKALPVPTEKTSKAKGQPEPOVYTLPPSREMTKN 445  
DB 243 VVSVLTVLHODWLNKGYCKVSNKALPVPTEKTSKAKGQPEPOVYTLPPSREMTKN 302  
QY 446 QVSLTCLVKGFPSPDIKAVEMESNGQPENNYKTTTPVLDSDGSFFLYSLKLTVDKSRQOQGVNF 505  
DB 303 QVSLTCLVKGFPSPDIKAVEMESNGQPENNYKTTTPVLDSDGSFFLYSLKLTVDKSRQOQGVNF 362  
QY 506 VFSCSVMEALHNYTKSLSLSPGK 531  
DB 363 VFSCSVMEALHNYTKSLSLSPGK 388  
RESULT 8  
US-08-897-236-23  
Sequence 23, Application US/08897236A  
Patent No. 6075007  
GENERAL INFORMATION:  
APPLICANT: Regeneron Pharmaceuticals, Inc.  
TITLE OF INVENTION: Modified Dorsal Tissue Affecting Factor and Composition  
FILE REFERENCE: REG 133  
CURRENT APPLICATION NUMBER: US/08/897,236A  
CURRENT FILING DATE: 1997-07-17  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 23  
LENGTH: 449  
TYPE: PRT  
ORGANISM: human  
US-08-897-236-23  
Query Match 43.3%; Score 1230; DB 3; Length 449;  
Best Local Similarity 56.1%; Pred. No. 2.8e-102;  
Matches 273; Conservative 21; Mismatches 95; Indels 98; Gaps 13;  
QY 105 LQSLPLLGQTLPALTVLDVSNRLTSLPLGALRGU-----139  
DB 1 MERCPSLGVTLVYLVV---LGLRATPAGQGHYLRHAPSDNLPVLVDLIEHPDIPDP 56  
QY 140 --GELQELYLKGNELKTLPPGLLTPTPKLEKLSLANNLTLPAGLLNGLNLTLLQOE 197  
DB 57 KEKDLNETLRLSLGCHYDPCFMATSPEDPRPGGGG-----AAGGAEDLAEQLDLQR 111  
QY 198 NSLYTIKPGFFGSHLLPFAFLHG-----NPWLCN---CEILYFRRLQDANAEVYVWKQV 249  
DB 112 -----PSGAMPSEIKGLEPSEGLAQGLQWLWSQTFPCVLY--AW---NDLGRFMPRY 160  
QY 250 VDV-----KAVTSNVASVQCDNSDKFPVYKYPCKGCTLGDEGDT 289  
DB 161 VKGSCFSKRSVPEGMVCKPSKSVHLTVLRWCQ-----RRGQRCGMPIQ---209  
QY 290 DLYDYTP-----BEDTEGDKVYPHTCPPEALGAPSVLPFPKPKDTLMISRTPEVT 344  
DB 210 -----YPIISCKSCSGDKT--HTCPPCPAPELGGSPVLPFPKPKDTLMISRTPEVT 262  
QY 345 CVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVSVLTVLHODWLNKGYCK 404  
DB 263 CVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVSVLTVLHODWLNKGYCK 322



QY 405 CKVSNKALPVPIEKTISKAKQPREPOVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVE 464  
Db 323 CKVSNKALPAPIEKTISKAKQPREPOVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVE 382  
QY 465 WESNGQPNENYKTTTPVLDSDGSGFFLYSKLTVDKSRWQOGNVFSCSVMEALHNNHYTQKS 524  
Db 383 WESNGQPNENYKTTTPVLDSDGSGFFLYSKLTVDKSRWQOGNVFSCSVMEALHNNHYTQKS 442  
QY 525 LSLSPGK 531  
Db 443 LSLSPGK 449

RESULT 9  
US-09-500-253B-23  
; Sequence 23, Application US/09500253B  
; Patent No. 6500640  
; GENERAL INFORMATION:  
; APPLICANT: Regeneron Pharmaceuticals, Inc.  
; TITLE OF INVENTION: Modified Dorsal Tissue Affecting Factor and Composition  
; FILE REFERENCE: REG 133-2  
; CURRENT APPLICATION NUMBER: US/09/500.253B  
; CURRENT FILING DATE: 2000-02-08  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 23  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: Human  
US-09-500-253B-23

Query Match 43.3%; Score 1230; DB 4; Length 449;  
Best Local Similarity 56.1%; Pred. No. 2.8e-102;  
Matches 273; Conservative 21; Mismatches 95; Indels 98; Gaps 13;  
QY 105 LQSLPLLGQTLPALTVLDVSNFRLTSLPLGALRGL----- 139  
Db 1 MERCPSLGVTLALVVV----LGLRATPAGOHVLTIRAPSDNLPVLDLIEHPDPIFP 56  
QY 140 --GELQELYKGNELKTLPPGLLTPTPKLEKLSLANNLTPALPAGLNGLENLDTLLQ 197  
Db 57 KEDLNELTLRLSLGCHYDPGFMTSPEDRPGGGG-----AAGGAEDLAELDQLLR 111  
QY 198 NSLYTIKPGFGSHLLPFAFLHG-----NPLCN---CEILYFRRWLQDIAENVYWKQV 249  
Db 112 -----PSGAMPSEIKGLEFSEGLAQGLQMLWMSQTFCPLY--AW----NDLGSREWP 160  
QY 250 VDV-----KAVTSNVASVQCDNSDKFPVYKYPGKGCPTIGDEGT 289  
Db 161 VKVGSFCFSKRSVPEGMVCKPSKVHLTVLRWCQ-----RRGGRCGWPIQ--- 209  
QY 290 DLYDYP-----EDTSGDKVVRHTCPCPAPEALGAPSVFLPPPKDPTLMISRTP 344  
Db 210 -----YPIIIECKKSCSGDKT--HTCPCPAPEALGAPSVFLPPPKDPTLMISRTP 262  
QY 345 CVVVDVSHEDPEVKFNWVDCGVEVHNAKTKPREQYNSTVRVSVLTVLHODWLNK 404  
Db 263 CVVVDVSHEDPEVKFNWVDCGVEVHNAKTKPREQYNSTVRVSVLTVLHODWLNK 382  
QY 405 CKVSNKALPVPIEKTISKAKQPREPOVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVE 464  
Db 323 CKVSNKALPAPIEKTISKAKQPREPOVYTLPPSREEMTKNOVSLTCLVKGFYPSDIAVE 382  
QY 465 WESNGQPNENYKTTTPVLDSDGSGFFLYSKLTVDKSRWQOGNVFSCSVMEALHNNHYTQKS 524  
Db 383 WESNGQPNENYKTTTPVLDSDGSGFFLYSKLTVDKSRWQOGNVFSCSVMEALHNNHYTQKS 442  
QY 525 LSLSPGK 531  
Db 443 LSLSPGK 449

RESULT 11

RESULT 10  
US-08-887-352B-14  
; Sequence 14, Application US/08887352B  
; Patent No. 5994511  
; GENERAL INFORMATION:  
; APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe  
; TITLE OF INVENTION: Improved Anti-IGE Antibodies and Method of  
; TITLE OF INVENTION: Improving Polypeptides  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/887.352B  
; FILING DATE: 03-Jul-1997  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Svoboda, Craig G.  
; REGISTRATION NUMBER: 39,044  
; REFERENCE/DOCKET NUMBER: P1123  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1489  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 451 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-887-352B-14

Query Match 43.1%; Score 1223; DB 2; Length 451;  
Best Local Similarity 64.1%; Pred. No. 1.2e-101;  
Matches 254; Conservative 14; Mismatches 58; Indels 70; Gaps 10;  
QY 188 ENLDTLLQENSLYTIKPGFF---GSHLLPFAFLHGNPMLCNCILYFRRWLQDIAENV 243  
Db 74 DSKNTFYQMLNSLRADTAVYTCARGSHY---FGH---W-----HFVWVGQGLTVTV 119  
QY 244 -----YVKKQVVDVKAVTSNVASVQCDNSDKFP-----VYKYPGKGC 280  
Db 120 SSASTKGPSVFLPAPSSKSTSGTAALGCLVKDYFPFVTVSWNSGALTSVHTFPA--- 176  
QY 281 PTLGDEG-----DTOLY-----DYPEBDTEGDKVRP-----HTCPCPA 315  
Db 177 -VLQSSGLYSLSSVTVFSSSLGTQYICNVNHKPSNTKYDKCKVPSKCDKTHTCPCPA 235  
QY 316 PEALGAPSVFLPPPKDPTLMISRTPVTCVVVDVSHEDPEVKFNWYDGVGEVHNAKTKP 375  
Db 236 PELLGAPSVFLPPPKDPTLMISRTPVTCVVVDVSHEDPEVKFNWYDGVGEVHNAKTKP 295  
QY 376 REEQYNSTVRVSVLTVLHODWLNKGYCKYCKSNKALPVPIEKTISKAKQPREQVYTL 435  
Db 296 REEQYNSTVRVSVLTVLHODWLNKGYCKYCKSNKALPVPIEKTISKAKQPREQVYTL 355  
QY 436 PPSREMTNQVSLTCLVKGFYPSDIAVEWESNGQPNENYKTTTPVLDSDGSGFFLYSKLT 495  
Db 356 PPSREMTNQVSLTCLVKGFYPSDIAVEWESNGQPNENYKTTTPVLDSDGSGFFLYSKLT 415  
QY 496 VDKSRWQOGNVFSCSVMEALHNNHYTQKLSLSPGK 531  
Db 416 VDKSRWQOGNVFSCSVMEALHNNHYTQKLSLSPGK 451



Patent No. 6037453  
 GENERAL INFORMATION:  
 APPLICANT: Jardieu, Paula M.  
 TITLE OF INVENTION: Immunoglobulin Variants  
 NUMBER OF SEQUENCES: 65  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genentech, Inc.  
 STREET: 1 DNA Way  
 CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: WinPatIn (Genentech)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/466,151  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/466163  
 FILING DATE: 06-Jun-1995  
 APPLICATION NUMBER: 08/405617  
 FILING DATE: 15-MAR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/185899  
 FILING DATE: 26-JAN-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/879495  
 FILING DATE: 07-MAY-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/744768  
 FILING DATE: 14-AUG-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Svoboda, Craig G.  
 REGISTRATION NUMBER: 39,044  
 REFERENCE/DOCKET NUMBER: P0718P2C1D1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650/225-1489  
 TELEFAX: 650/952-9881  
 INFORMATION FOR SEQ ID NO: 65:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 451 amino acids  
 TYPE: Amino Acid  
 TOPOLOGY: Linear  
 US-08-466-151-65

Query Match 43.1%; Score 1223; DB 3; Length 451;  
 Best Local Similarity 64.1%; Pred. No. 1.2e-101;  
 Matches 254; Conservative 14; Mismatches 58; Indels 70; Gaps 10;  
 QY 188 ENLDTLQLQNSLYTIPKGF---GSHLLPFAFLHGNPWLNCNCEILYFRRLQDANV 243  
 DB 74 DSKNTFYLQNSLRAEDTAVYICARGSHY---FGH---W-----HFAVWGQGLTVTV 119  
 QY 244 -----YVMQVVDVKAVTSNVAQVQDNDKFP-----VYKYPGKGC 280  
 DB 120 SSASTKGPSVFPLAPSSKSTSGGTAALGLVQDYFPFPTVTSWNSGALTSGVHTFPA--- 176  
 QY 281 PTLDGEG-----DLDLY-----DYYPEDETEGDKVRP-----HTCPCPA 315  
 DB 177 -VLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHNKPSNTKVKDKEPKSCDKTHTCPCPA 235  
 QY 316 PEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKTP 375  
 DB 236 PELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKTP 295  
 QY 376 REQYNSTRYVSVTLVTHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTL 435  
 DB 296 REEQYNSTRYVSVTLVTHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTL 355  
 QY 436 PSREEMTKNQVSLTCLVKGFYPSDIAVEGSGNQENNNYKTTTPVLDSDGSFFLYSKLT 495  
 DB 356 PPSREEMTKNQVSLTCLVKGFYPSDIAVEGSGNQENNNYKTTTPVLDSDGSFFLYSKLT 415  
 QY 496 VDKSRWQOGNVFSCSVNHEALHNHYTQKSLSLSPGK 531  
 DB 416 VDKSRWQOGNVFSCSVNHEALHNHYTQKSLSLSPGK 451

RESULT 15  
 US-09-109-207C-16  
 Sequence 16, Application US/09109207C  
 Patent No. 6172213  
 GENERAL INFORMATION:  
 APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe  
 TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of Improving Polypept

QY 436 PPSREEMTKNQVSLTCLVKGFYPSDIAVEGSGNQENNNYKTTTPVLDSDGSFFLYSKLT 495  
 DB 356 PPSREEMTKNQVSLTCLVKGFYPSDIAVEGSGNQENNNYKTTTPVLDSDGSFFLYSKLT 415  
 QY 496 VDKSRWQOGNVFSCSVNHEALHNHYTQKSLSLSPGK 531  
 DB 416 VDKSRWQOGNVFSCSVNHEALHNHYTQKSLSLSPGK 451  
 RESULT 14  
 US-09-109-207C-14  
 Sequence 14, Application US/09109207C  
 Patent No. 6172213  
 GENERAL INFORMATION:  
 APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe  
 TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of Improving Polypept  
 FILE REFERENCE: P1123R1  
 CURRENT APPLICATION NUMBER: US/09/109,207C  
 CURRENT FILING DATE: 1998-06-30  
 PRIOR APPLICATION NUMBER: US 60/051,554  
 PRIOR FILING DATE: 1997-07-03  
 NUMBER OF SEQ ID NOS: 44  
 SEQ ID NO 14  
 LENGTH: 451  
 TYPE: PRT  
 ORGANISM: Artificial  
 FEATURE:  
 NAME/KEY: Artificial  
 LOCATION: 1-451  
 OTHER INFORMATION: Heavy chain sequence derived from MAE11  
 US-09-109-207C-14

Query Match 43.1%; Score 1223; DB 3; Length 451;  
 Best Local Similarity 64.1%; Pred. No. 1.2e-101;  
 Matches 254; Conservative 14; Mismatches 58; Indels 70; Gaps 10;  
 QY 188 ENLDTLQLQNSLYTIPKGF---GSHLLPFAFLHGNPWLNCNCEILYFRRLQDANV 243  
 DB 74 DSKNTFYLQNSLRAEDTAVYICARGSHY---FGH---W-----HFAVWGQGLTVTV 119  
 QY 244 -----YVMQVVDVKAVTSNVAQVQDNDKFP-----VYKYPGKGC 280  
 DB 120 SSASTKGPSVFPLAPSSKSTSGGTAALGLVQDYFPFPTVTSWNSGALTSGVHTFPA--- 176  
 QY 281 PTLDGEG-----DLDLY-----DYYPEDETEGDKVRP-----HTCPCPA 315  
 DB 177 -VLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHNKPSNTKVKDKEPKSCDKTHTCPCPA 235  
 QY 316 PEALGAPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKTP 375  
 DB 236 PELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKTP 295  
 QY 376 REQYNSTRYVSVTLVTHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTL 435  
 DB 296 REQYNSTRYVSVTLVTHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTL 355  
 QY 436 PSREEMTKNQVSLTCLVKGFYPSDIAVEGSGNQENNNYKTTTPVLDSDGSFFLYSKLT 495  
 DB 356 PPSREEMTKNQVSLTCLVKGFYPSDIAVEGSGNQENNNYKTTTPVLDSDGSFFLYSKLT 415  
 QY 496 VDKSRWQOGNVFSCSVNHEALHNHYTQKSLSLSPGK 531  
 DB 416 VDKSRWQOGNVFSCSVNHEALHNHYTQKSLSLSPGK 451

RESULT 15  
 US-09-109-207C-16  
 Sequence 16, Application US/09109207C  
 Patent No. 6172213  
 GENERAL INFORMATION:  
 APPLICANT: Henry B. Lowman, Leonard G. Presta, Paula M. Jardieu, John Lowe  
 TITLE OF INVENTION: Improved Anti-IgE Antibodies and Method of Improving Polypept

